

PhD Thesis

**Mobile Thermography-based
Physiological Computing
for Automatic Recognition of
a Person's Mental Stress**

Youngjun Cho

January 2019

UCL Interaction Centre

UCL

I, Youngjun Cho, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signature:

Abstract

This thesis explores the use of *Mobile Thermography*¹, a significantly less investigated sensing capability, with the aim of reliably extracting a person's multiple physiological signatures and recognising mental stress in an automatic, contactless manner. Mobile thermography has greater potentials for real-world applications because of its light-weight, low computation-cost characteristics. In addition, thermography itself does not necessarily require the sensors to be worn directly on the skin. It raises less privacy concerns and is less sensitive to ambient lighting conditions.

The work presented in this thesis is structured through a three-stage approach that aims to address the following challenges: i) thermal image processing for mobile thermography in variable thermal range scenes; ii) creation of rich and robust physiology measurements; and iii) automated stress recognition based on such measurements. Through the first stage (Chapter 4), this thesis contributes new processing techniques to address negative effects of environmental temperature changes upon automatic tracking of regions-of-interest and measuring of surface temperature patterns. In the second stage (Chapters 5,6,7), the main contributions are: robustness in tracking respiratory and cardiovascular thermal signatures both in constrained and unconstrained settings (e.g. respiration: strong correlation with ground truth, $r=0.9987$), and investigation of novel cortical thermal signatures associated with mental stress. The final stage (Chapters 8,9) contributes automatic stress inference systems that focus on capturing richer dynamic information of physiological variability: firstly, a novel respiration representation-based system (which has achieved state-of-the-art performance: 84.59% accuracy, two stress levels), and secondly, a novel cardiovascular representation-based system using short-term measurements of nasal thermal variability and heartrate variability from another sensing channel (78.33% accuracy achieved from 20seconds measurements). Finally, this thesis contributes software libraries and incrementally built labelled datasets of thermal images in both constrained and everyday ubiquitous settings. These are used to evaluate performance of our proposed computational methods across the three-stages.

¹ *Thermography* is also called thermal imaging. In this thesis, *Mobile Thermography* is defined as thermal imaging using mobile, low-cost thermographic systems allowing them to perform in unconstrained, mobile, real world situations.

Impact Statement

This thesis proposes a new pathway to physiological and stress measurements through mobile thermal imaging. State-of-the-art performance achieved from tests with the use of small, light-weight, and more importantly, low-cost sensing devices makes itself more feasible and practical in supporting real-world applications. Its physiological computing² capability alone can benefit general healthcare and fitness sectors by providing contactless, simultaneous measurements of multiple vital signs to specific user groups, for example, people with chronic pain who tend to reduce the number of clothes and/or devices that touch their body. The physiological computing-empowered mental stress detection capability can produce fundamental beneficiaries to this modern society where mental stress has become a critical problem.

This form of stress strongly affects our wellbeing. With disabled people for instance, even basic everyday tasks such as using public transport to go shopping can induce high level stress. However, people are often not aware of their stress level until it becomes too high. Long exposure to this situation leads to further impairment of their mental capability, such as attention, memory and decision-making (Yerkes & Dodson, 1908). The ability to automatically recognise a person's mental stress is fundamental to the personalisation of better and continuous stress management support. To date however, the only technologies which are available for this are intrusive, heavy, fragile, or restrict mobility and movement, therefore limiting their use. On the other hand, our proposed advanced tracking methods which enable low-cost, mobile thermography to have the ability to create new opportunities for a wide range of user groups, spanning from therapists to the general public.

This thesis has also strong industrial impact. Low-cost mobile thermal cameras first arrived on the market a couple of years ago. However, tools for mobile thermal imaging have remained very limited. Our proposed novel frameworks and computational methods, which activate reliable mobile thermal imaging, and furthermore physiology/stress sensing in

² In this thesis, we define *physiological computing* as technology that listens to a person's bodily (physiological and psychological) states and needs.

unconstrained real-world settings, can help build new industrial areas. Indeed, we have applied our proposed approach to industrial sectors through the EU H2020 Human research project³ in order to help tailor factory workers' work schedule towards their psychological needs. We have also tested our own distinctive advantages against commercially available sensors (e.g. bracelet-based devices interfering with machine operation). Already, some industrial partners (i.e. small and big companies⁴) have expressed their interest with building consumer solutions.

Finally, research-driven software libraries, datasets and tools made within this thesis, which have been released to foster this research, can have a strong impact upon this area of research. Many academic and industrial research institutions have had access to the released outcomes. For example, our released datasets have been used by circa 20 research institutions (by October 2018)⁵. All in all, this thesis is expected to contribute towards helping shape our lives in beneficial ways.

³ <http://www.humanmanufacturing.eu/>

⁴ For example, Samsung, LG, Sony and Huawei, etc.

⁵ Another example is our paper published at CHI'18 (Cho *et al.*, 2018), which has been downloaded more than 800 times within 6 months (between late April and mid-October in 2018).

Acknowledgements

First of all, I would like to express my sincere, special thanks to Nadia Bianchi-Berthouze who is my primary supervisor and mentor. She has provided excellent support and supervision to me throughout all stages of preparation of this thesis. I have always wanted to say “you are one of the most outstanding, enthusiastic scholars I have ever known”. I am also immensely grateful to my secondary supervisors, Nicolai Marquardt and Simon J. Julier for giving me invaluable feedback and advice. All their support and inputs have helped produce this work of which I am very proud. To acknowledge this, I have used the first-person plural (“*We*”) rather than “*I*” through my thesis.

I would like to sincerely thank PhD examiners of this thesis, Mirco Musolesi (internal: UCL, Alan Turing Institute), Albrecht Schmidt (external: Ludwig Maximilian University of Munich) and my upgrade viva examiner, Anthony Steed for providing me with their insightful feedback and advice.

I have been very fortunate to have great research, social and interactive environments provided by UCL Interaction Centre (UCLIC). I owe many thanks to all UCLIC academics for their supports, in particular, Yvonne Rogers, Catherine Holloway and Duncan Brumby. I am also very grateful to Louise Gaynor for her useful help over the years. I also wish to thank all the members of Nadia’s affective computing group. I would like to thank all UCLIC researchers and PhD students.

I wish to acknowledge that this research has been funded by University College London Overseas Research Scholarship (UCL-ORS) which is designed to support top quality international postgraduate students, and this work has been additionally supported by the EU H2020 Human studentship. I would like to thank their support.

I owe many thanks to all my old friends and my previous colleagues who I worked with at LG and KAIST many years ago.

I would like to express my very grateful thanks to Gabriela Gordon who has energised me, proofread draft material and has been hugely supportive.

I would like to express my deepest gratitude to my parents and sister for encouraging and teaching me to realise the most important things in life. Without all their support and love, this thesis would not have existed.

Thank you very much!

Selected Publication List related to this thesis

* **Youngjun Cho**, Simon J. Julier, Nicolai Marquardt and Nadia Bianchi-Berthouze, 2017. Robust Tracking of Respiratory Rate in High-Dynamic Range Scenes using Mobile Thermal Imaging. *Biomedical Optics Express* 8(10):4480–4503.

(Research article, Chapters 4 and 5)

* **Youngjun Cho**, Nadia Bianchi-Berthouze, Nicolai Marquardt and Simon J. Julier, 2018. Deep Thermal Imaging: Proximate Material Type Recognition in the Wild through Deep Learning of Spatial Surface Temperature Patterns. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI'18)*: 2:1–2:13.

(Research article, Chapter 4)

* **Youngjun Cho**, Simon J. Julier and Nadia Bianchi-Berthouze, 2017. DeepBreath: Deep Learning of Breathing Patterns for Automatic Stress Recognition using Low-Cost Thermal Imaging. *The 7th International Conference on Affective Computing and Intelligent Interaction 2017 (ACII'17)*: 456–463.

(Research article, Chapter 8)

* **Youngjun Cho**, Simon J. Julier and Nadia Bianchi-Berthouze, 2019. Instant Stress: Detection of Perceived Mental Stress through Smartphone Photoplethysmography and Thermal Imaging. *Journal of Medical Internet Research / JMIR Mental Health - Special Issue on Computing and Mental Health (2018)* (in press).

(Research article, Chapter 9)

Selected Publication List unrelated, but during my PhD study

* **Youngjun Cho**, 2017. Sensorless Resonance Tracking of Resonant Electromagnetic Actuator through Back-EMF Estimation for Mobile Devices, arXiv:1803.07065.

(Research article)

* **Youngjun Cho**, Andrea Bianchi, Nicolai Marquardt, and Nadia Bianchi-Berthouze. 2016. RealPen: Providing Realism in Handwriting Tasks on Touch Surfaces using Auditory-Tactile Feedback. In Proceedings of the 29th Annual Symposium on User Interface Software and Technology (UIST'16): 195-205.

(Research article)

Full publication list available online: <http://scholar.google.com/citations?user=2h-v48AAAAAJ>

Contents

1	Introduction.....	35
1.1.	Background	35
1.2.	Research Questions, Scopes and Contributions	38
1.3.	Thesis Structure	45
2	Background	47
2.1.	An Overview: from earlier Medical Thermometry to state-of-the-art Thermal Imaging.....	47
2.2.	Thermography as a Measure of Physiological Cues	51
2.3.	Thermography as a Measure of Human Affects.....	56
2.3.1.	Thermal Directionality in relation to a Person’s Affective States.....	58
2.3.2.	Facial Muscular Thermal Signature in relation to a Person’s Affective States	63
2.3.3.	Summary Tables	64
2.4.	The Scalp as a Potential Region of Interest (ROI) for Stress-related Thermal Signatures: Neuroscientific Evidence.....	69
2.4.1.	Stress Responses of the Brain in relation to Temperature.....	70
2.4.2.	Potential Local Scalp Regions providing Cortical Thermal Signatures....	71
2.5.	Chapter Summary and Problem Statements	74
3	Research Methodology	79
3.1.	Three Stage Approach	79
3.2.	Basic Tools for Thermal Imaging in Unconstrained Settings	85
3.3.	Overview of Experimental Protocols and Datasets	87
3.4.	Overview of Evaluation Methods.....	93
3.5.	Chapter Summary	94
	STAGE I.....	97
4	Optimal Quantisation: A Preprocessing method for Mobile Thermography in Variable Thermal Range Scenes.....	99
4.1.	Background and Research Questions	99
4.2.	Proposed Computational Methods	102
4.2.1.	Optimal Quantisation Algorithm.....	102

4.2.2. Optimal Quantisation-powered Automatic ROI Tracking.....	103
4.3. Experimental Protocol and Dataset (<i>ROITracking</i>).....	105
4.4. Evaluation Method.....	107
4.5. Results.....	107
4.6. Generalisation to Other Material Temperatures.....	109
4.6.1. Method: Deep Thermal Imaging	109
4.6.2. Data Collection Study.....	113
4.6.3. Results	116
4.7. Discussion.....	119
4.8. Chapter Summary	121
STAGE II	125
5 Robust Respiration Tracking in Unconstrained Settings	127
5.1. Background and Research Question	127
5.2. Proposed Computational Methods.....	131
5.2.1. Robust Nostril-ROI Tracking: <i>Thermal Gradient Flow</i>	132
5.2.2. Respiration Estimation by <i>Thermal Voxel</i> Integration	134
5.3. Experimental Protocols and Datasets (<i>ThermalBreath I, II</i>).....	137
5.4. Evaluation Method.....	140
5.5. Results.....	143
5.5.1. Nostril-region Tracking Performance.....	143
5.5.2. Respiration Tracking Performance	146
5.6. Discussion	151
5.7. Chapter Summary	155
6 Robust Tracking of Stress-induced Nasal Thermal Variability	157
6.1. Background and Research Question	157
6.2. Proposed Computational Methods.....	160
6.2.1. Continuous Tracking of Nasal Temperature.....	160
6.2.2. Metrics for Quantification of Nasal Thermal Variability	163
6.3. Experimental Protocols and Datasets (<i>NoseTracking, StressNose</i>)	165
6.4. Evaluation Method.....	168
6.5. Results.....	169
6.5.1. Comparison between Small and Large ROIs.....	169
6.5.2. Effect of Stressors on Thermal Variability Metrics.....	172
6.6. Discussion	174
6.7. Chapter Summary	176
7 Exploring Stress-induced Cortical Thermal Signatures from the Scalp	179
7.1. Background and Research Question	180
7.2. Study I: Initial Exploration in Unconstrained Setting.....	184
7.2.1. Recording Set-up	185
7.2.2. Experimental Protocol and Dataset (<i>ThermalBrain I</i>).....	185

7.2.3. Evaluation Method.....	187
7.2.4. Results.....	188
7.2.5. Reflection on Study I: Identified Challenges	191
7.3. Study II: London Arithmetic Stress Task (LAST) in Constrained Setting.....	192
7.3.1. Recording Set-up.....	192
7.3.2. Proposed Computational Methods	194
7.3.3. Experimental Protocol and Dataset (<i>ThermalBrain II</i>)	198
7.3.4. Evaluation Method.....	201
7.3.5. Results.....	202
7.4. Discussion	206
7.5. Chapter Summary	210
STAGE III	213
8 Automated Mental Stress Recognition using Respiratory Signature.....	215
8.1. Background and Research Question.....	216
8.2. Proposed Computational Methods	218
8.2.1. Novel Representation of Respiration Variability	218
8.2.2. Automatic Feature Learning using Convolutional Neural Network (CNN)	220
8.3. Experimental Protocol and Dataset (<i>DeepBreath</i>)	222
8.4. Evaluation Method	224
8.5. Perceived Mental Stress Levels and Labelling.....	225
8.6. Results	227
8.7. Discussion	230
8.8. Chapter Summary	233
9 Instant Automated Inference of Mental Stress using Cardiovascular Signatures	235
9.1. Background and Research Question.....	235
9.2. Recording Set-up	238
9.3. Proposed Computational Method	240
9.4. Experimental Protocol and Dataset (<i>InstantStress</i>).....	241
9.5. Evaluation Method	243
9.6. Perceived Mental Stress Levels and Labelling.....	244
9.7. Results	245
9.8. Discussion	252
9.9. Chapter Summary	257
10 Conclusions.....	259
10.1. Mobile Thermal Imaging: Bringing Thermal Imaging into Unconstrained, Mobile, Real-world Settings.....	262
10.2. ROI Tracking and Environmental Temperature Changes	266

10.3. Representation of Physiological Thermal Signature	267
10.4. Important Role of Tracking Physiological Variability in Stress Assessment	268
10.5. Limitation and Future Directions	270
References	273
Appendices	293
<i>Appendix A. (Extra Study) Skeletal Muscle Activation as a New Thermal Signature from Underexplored ROIs: Understanding of Thermal Signatures of Physical Stress during Anaerobic Exercise.....</i>	<i>293</i>
A.1. Background: Physical Stress and Muscle Temperature.....	293
A.2. Experimental Protocol and Dataset.....	296
A.3. Results	299
<i>Appendix B. Extraction of Cardiac Pulse Rate.....</i>	<i>304</i>
<i>Appendix C. Extraction of Eye Blinks</i>	<i>306</i>
<i>Appendix D. Exploration of Non-linearity in Quantisation Process</i>	<i>307</i>
D.1 Method.....	307
D.2. Results	308
<i>Appendix E. Additional Materials for Deep Thermal Imaging (Material Type Recognition).....</i>	<i>310</i>
<i>Appendix F. BVP and P-P Interval Estimation through PPG Imaging</i>	<i>312</i>
F.1. Spatial Entropy-based camera PPG.....	312
F.2. Result of BVP Signal Quality Test.....	313
<i>Appendix G. Hand-Engineered Features for Thermal Analysis.....</i>	<i>315</i>
<i>Appendix H. Questionnaires.....</i>	<i>316</i>
<i>Appendix I. Ethical Approval.....</i>	<i>319</i>

Glossary

Terminology	Definition
Autonomic Nervous System (ANS)	the part of the nervous system responsible for control of the bodily functions not consciously directed, such as breathing, the heartbeat, and digestive processes.
Cardiovascular	relating to the heart and blood vessels.
Cerebral Blood Flow (CBF)	the blood supply to local functional areas of the brain.
Convolutional Neural Network (CNN)	an artificial neural network built mainly with convolutional layers.
Cortical	relating to the outer layer of the cerebrum.
Corticotropin-Releasing Hormone (CRH)	a peptide hormone or neurotransmitter responding to stressors. Also known as corticotropin-releasing factor (CRF).
Directionality of temperature	a characteristic in terms of a binary direction of thermal variations.
Stress	[definition 1] (Selye, 1980) the nonspecific response of the body to any demand. [definition 2] (Oxford English Dictionary, 2018) a state of mental or emotional strain or tension resulting from adverse or demanding circumstances.
Spectrogram	a photographic or other visual or electronic representation of a spectrum.
Mobile Thermography	thermal imaging using a mobile, low-cost thermographic system allowing themselves to perform in unconstrained, mobile, real world situations. It is also called mobile thermal imaging.
Nasal	relating to the nose.
Perspiratory	relating to or affecting sweating or the sweat glands.
Physiological Computing	technology that listens to a person's bodily (physiological and psychological) states and needs (definition proposed in the thesis).
Quantisation	a process to approximate (a continuously varying signal) by one whose amplitude is restricted to a prescribed set of values.
Region of Interest (ROI)	a selected local region of an image.
Respiratory	relating to or affecting respiration (breathing) or the organs of respiration.
ROI Tracking	a method to keep track of the position of an region of interest (ROI) on image sequences.
Skew	(of a statistical distribution) not symmetrical.
Skewness	a measure of the asymmetry of a probability distribution

Thermal	relating to heat.
Thermography	a method to measure, and visualise, spatial temperature distributions of (an) object(s), materials and scenes, etc. It is also called thermal imaging.
Variability	lack of consistency or fixed pattern, liability to change.
Vasoconstriction	the narrowing of blood vessels.
Vasodilation	the widening of blood vessels

List of Figures

Figure 1.1. (a) an example of a thermal image of the author taken by (b) a high-precision, high-cost, heavyweight thermal camera (FLIR SC5000mb) – dimension: 320x141x159 mm ³ ; weight: 3,800.0g (this camera also needs to be used with a powerful desktop/laptop for data collection). And (c) a low-precision, low-cost, lightweight thermal camera (FLIR One 2G) – dimension: 34x68x14 mm ³ ; weight: 36.5g.....	37
Figure 1.2. Research Question (RQ) Tree: the main research question leads to detailed sub-questions along with three themes: i) signal processing techniques for mobile thermal imaging, ii) physiology measurements through mobile thermal imaging, iii) automated stress recognition based upon mobile thermal imaging	39
Figure 2.1. A widely used experimental setup in the literature (adapted from Pavlidis <i>et al.</i> 2012).....	54
Figure 2.2. A summary of the literature exploring the use of thermography as a measure of physiological cues and affects: choosing right ROIs corresponding to physiological signatures of interest and their metrics are key in assessing a person's affective states (Thickness of a line is to highlight what a majority of existing works have focused on). All existing works have been conducted in systematic constrained settings in terms of environmental temperature changes and movements.....	57
Figure 2.3. Left) an example of thermal image of his face and thumb finger, and widely used ROIs, Right) a quick summary of thermal directional responses of selected ROIs to a person's affective states.	58
Figure 3.1. Mobile thermal imaging as a new way for bringing thermal imaging from traditional constrained systematic settings into unconstrained, mobile, real-world like situations.....	79
Figure 3.2. A low-cost thermal camera (FLIR One) to be used as a mobile thermal imaging channel in the thesis. This device requires a smartphone to be connected to control parameters for itself and to visualise, and store, thermal images. They can be either directly connected or connected with an extension cable.	86
Figure 3.3. (a) Thermal image with bicubic interpolation and scaling, (b) the proposed data frame protocol.....	87
Figure 3.4. 10cm VAS based questionnaire with examples of labels along with each session type (R ₁ , R ₂ : Rest from Session 1 and 2, Se: Stroop easy, Sh: Stroop hard, Me: Math easy, Mh: Math hard). The red marks (x) represent examples of self-reported scores of one participant over the different tasks. The task labels have been added by the researchers for the purpose of this figure.	91

Figure 4.1. Key challenge in the ROI tracking through thermal imaging in high thermal dynamic range scenes: fixed thermal range of interest is not suitable in preserving the morphological facial shape within varying ambient temperature: [top] examples of thermogram shots collected from a person walking outdoor (for 6 minutes), [bottom] temperature histograms.	101
Figure 4.2. The proposed optimal quantisation and a processing flow for automatic tracking of a region of interest (ROI) and the extractions of thermal variances from the ROI.	104
Figure 4.3. Experimental setup: a low-cost thermal camera was installed near each participant's face using a shoulder rig. An office room was used to simulate everyday desk activities. Participants were not asked to constrain their movements. Additionally, a chest-belt type respiration sensor was used. The data is not used in this chapter but used as a reference signal in the following chapter where we evaluate respiration tracking algorithm performance.	105
Figure 4.4. Examples of collected facial thermal images of participants. Not constraining participants' movements helped to capture their very natural motions such as lowering, turning, and rotating their heads.	106
Figure 4.5. The overall process for proximate material type recognition with Deep Thermal Imaging: a) physical sensing via mobile thermal imaging, b) dynamic range quantisation on every thermal image frame, c) feature learning and training process based on spatial transformer and deep convolutional neural networks, d) prediction results.	110
Figure 4.6. Implementation of Deep Thermal Imaging, the real-time proximate material type recogniser.	113
Figure 4.7. Outdoor materials collected for the material classification task. .	114
Figure 4.8. Data collection: thermal images of outdoor materials were recorded by two experimenters while walking/running outside both at daytime and night-time in different seasons.	115
Figure 4.9. Examples of collected and quantised thermal images of t. iron fence, +. flower bed, u. granite stones (pavement edge) and §. building entrance carpet. These images contain some degree noise, e.g., dirt and rubbish.....	115
Figure 4.10. Thermal dynamics levels for the total 26584 thermal images across all 17 outdoor materials (p-ç). The thermal dynamics were estimated by using $(T_2 - T_1)$ in Equation (4.5). The thermal dynamic range of indoor materials (in Appendix E) is marked with red dot lines (A) for comparison.....	116
Figure 4.11. Mean class accuracy results (17 classes) from i) the proposed method, but without the Dynamic Range Quantisation (DRQ) and ii) that with the DRQ (error bar: 95% confidence interval).	117
Figure 4.12. Mean class accuracy results using 5-fold cross validation: (a) without, (b) with the support of Dynamic Range Quantisation ...	117
Figure 4.13. 10-fold cross validation results (<i>high dynamic temperature range proximate material, labels for materials shown in Figure 4.7</i>): (a) confusion matrix (in percentage; colour range: white(0%)-	

	black(100%)), (b) mean class accuracy recognition distribution (error bar: 95% confidence interval).	118
Figure 5.1.	Key challenges in respiration tracking through thermal imaging in real-world settings: (a) difficulty in tracking the nostril ROI: the shape of the nostril is affected by one's mobility and thermal dynamics, (b) low signal quality of respiration patterns: four examples of thermal images show the tracked nostril region while breathing. The traditional average temperature is not ideal in extracting respiratory patterns when the respiration-driven thermal variance is weak, e.g. during shallow breathing (Case 2) compared with deep breathing (Case 1). The low spatial resolution of mobile thermal imaging also leads to the weak signal.	130
Figure 5.2.	(a) Overall procedure of thermal imaging-based respiration tracking, (b) key components: 1) Optimal Quantisation (proposed in Chapter 4), 2) Thermal Gradient Flow –nostril-region tracking method based on the thermal gradient magnitude computation and points tracking methods, 3) Thermal Voxel-based Respiration Estimation – extracting the respiratory signals by integrating the unit thermal-voxels inside the nostril.	131
Figure 5.3.	Example shots of conversion from thermal images to thermal-gradient magnitude maps: the proposed method can help to preserve the morphology of the nostril region during motion (Zoomed-in-areas are rotated for the visual representation).	132
Figure 5.4.	Extraction of respiratory patterns through Thermal Voxel integration: (a) a person's nostril and its thermogram sequences along the time in 3D (top) and 2D (bottom), (b) the concave volume corresponding to heat variances in the nostril, (c) the extracted respiratory signals compared with ground truth signals, and (d) a comparison of the filtered voxel-based signals with the traditional method as the participant changes their head. The voxel method closely tracks ground truth, but the traditional method fails.	135
Figure 5.5.	Experiment with controlled respiration in environments with non-constant temperature: (a) to obtain different thermal dynamic range scenes (i.e. environments with non-constant dynamic temperature), four different places were chosen (Place A: room, B: entrance of the building, C: corner on the street, D: park), the last image-shot is a thermal image collected in Place D, the experiment was run in winter, (b) the guiding breathing patterns are composed of three different rates (10(slow), 15(normal), 30(fast) breaths/min).	138
Figure 5.6.	Experiment with unconstrained respiration in both indoor and outdoor light physical activities.	139
Figure 5.7.	Statistical methods for evaluations: (a) automated-synchronization between estimated signals and reference signals using the <i>maximum-amplitude of cross-correlation (MACC)</i> , (b) respiration-related goodness probability as a respiratory signal quality index (SQI).	142
Figure 5.8.	Quantified motion artefacts using the relative Euclidean distance from the origin of the nostril-ROI at the first frame: (a) example from <i>ThermalBreath I</i> (fully controlled), (b) from <i>ROITracking</i>	

	(sedentary activity), and (c) from <i>ThermalBreath II</i> (physical activity).	144
Figure 5.9.	Overall results of the nostril-tracking performance of the proposed Thermal Gradient Flow (black) compared with existing methods: (a) <i>ThermalBreath I</i> (controlled but in non-constant temperature scenes), (b) <i>ROITracking</i> (unconstrained respiration during sedentary activity), (c) <i>ThermalBreath II</i> (unconstrained respiration during physical activity).....	144
Figure 5.10.	Results of respiratory signature extraction: (a,c,e) time-domain signals, (b,d,f) the respiration-related goodness metric values. The proposed Thermal Voxel-based method is more robust than the traditional averaging-based method for Situation 1 (a-b) and Situation 2 (c-d). For Situation 3 (e-f) (i.e. fully mobile contexts), the ground truth shows less reliability in the respiration tracking, except for the segment A in (e) (i.e. standing with small movement).	147
Figure 5.11.	Overall results from Situation 1 (<i>ThermalBreath I</i>): Bland-Altman plots of (a) Thermal Voxel integration method, (b) the traditional averaging method, and (c) overall RMSE comparisons.	148
Figure 5.12.	Separate results along with the different environment in Situation 1 (<i>ThermalBreath I</i>): (a) Bland-Altman plots of Thermal Voxel integration method and the traditional averaging method, and (b) RMSE comparisons for Place A – D.	149
Figure 5.13.	Situation 2 (<i>ROITracking</i>): Bland-Altman plots of (a) Thermal Voxel integration method, (b) the traditional averaging based estimation method, and (c) overall RMSE comparisons. The mean P_r value (pSQI) from the ground truth in Situation 1 (<i>ThermalBreath I</i>) was set as exclusion criterion ($P_r \geq 0.9825$). ..	150
Figure 5.14.	Situation 3 (<i>ThermalBreath II</i>): Bland-Altman plots of (a) Thermal Voxel integration method, (b) the traditional averaging method, and (c) overall RMSE comparisons. The mean P_r value (pSQI) from the ground truth in Situation 1 (<i>ThermalBreath I</i>) was set as exclusion criterion ($P_r \geq 0.9825$).	151
Figure 6.1.	Examples of thermal images of a person’s face and the nose tip selected as a ROI: the ROI from each thermal image has different temperature distribution, not providing common point features or patterns across thermal images which are required for the ROI tracking.....	158
Figure 6.2.	Selection of a larger ROI that includes the nose tip (small ROI) to obtain strong facial point features to enable reliable ROI tracking: (left) raw thermal image, (right) thermal gradient image. White stars (*) represent facial point features.	159
Figure 6.3.	The proposed method for tracking of the nose tip temperature variances: Step I) selecting a larger ROI and tracking, Step II) computing average of temperatures within the ROI, Step III) removing outliers with a sliding window, Step IV) removing high frequency components.	161
Figure 6.4.	A person’s respiratory activity influences the nasal tip temperature: (a) examples of thermal images (view angles were not constrained),	

	(b) the nasal temperature changes during inhalation (yellow: warmer, red: moderate, black: colder).	162
Figure 6.5.	Post-processing (down sampling and feature scaling using the min-max approach) for the extraction of one-dimensional feature-scaled thermal variability sequence from a person's nose tip.....	163
Figure 6.6.	Experimental setup (the image was taken by a thermal camera) for the data collection for Dataset I: a chin rest was used to maintain the position of the nose of each participant still. The small ROI (reference) includes only the nose tip and the large ROI contains the nose tip and its surrounding area (e.g. part of the nostril).....	166
Figure 6.7.	Flow chart of the study protocol (*Counterbalanced in Latin squared design).....	167
Figure 6.8.	Experimental setup: a smart phone with a low-cost thermal camera was installed and running on the desk while each person participated in the tasks. The participant was not asked to constrain their movements.	168
Figure 6.9.	An example of extracted average temperature and (temporally) filtered temperature sequences from a small ROI and a large ROI on facial thermal images of Participant 2.....	170
Figure 6.10.	Scatter plots of (a) average temperatures and (b) (temporally) filtered average temperatures from the chosen small ROI and the large ROI (from 10 participants, N=1000).....	170
Figure 6.11.	Bar plot of Pearson correlation coefficients between each pair of samples (both data from small and large ROIs) for every individual (10 participants, each coefficient was computed from each participant's samples (N=100)): (a) average temperature timeseries and (b) the filtered data.	171
Figure 6.12.	Box plots of respiratory signal quality test results about average temperatures from both small and large ROIs and the filtered average temperatures from the large ROI (the proposed signal to use) using the respiratory pSQL.....	172
Figure 6.13.	Boxplot (95% confidence interval) of the self-reported perceived mental stress scores of 12 participants across each task (Rest, Math Easy, Math Hard). * describes the significant difference. σ^7 (P7, Rest) is an statistical outlier.	172
Figure 7.1.	Location of the <i>left inferior frontal gyrus</i> (Red and yellow colour), the most commonly activated region under stress conditions (Taylor et al., 1997; Soufer et al., 1998; Pruessner et al., 2008): on the brain images synthesised using the Neurosynth platform (Yarkoni et al., 2011), we additionally draw facial landmarks for the purpose of visualisation. x,y,z values represent MNI coordinates and D: Dorsal, V: Ventral, A: Anterior, P: Posterior, L: Left, R: Right.	181
Figure 7.2.	Location of the <i>right prefrontal cortex</i> and the <i>right dorsolateral prefrontal cortex</i> , the most commonly activated region under stress conditions (Wang et al., 2005; Causse et al., 2017; Rosenbaum et al., 2018): we segment the brain into 3 (anterior, middle, posterior) x 2 (left, right) areas on the z -axis images. x,y,z values represent MNI coordinates and D: Dorsal, V: Ventral, A: Anterior, P: Posterior, L: Left, R: Right.	182

- Figure 7.3.** Location of the *parietal/occipital lobes* (the *occipital cortex* and *right inferior parietal lobules*), the frequently reported “stress” activation region (Taylor et al., 1997; Pruessner et al., 2008): we segment the brain into 3 (anterior, middle, posterior) x 2 (left, right) areas on the *z-axis* images. *x,y,z* values represent MNI coordinates and D: Dorsal, V: Ventral, A: Anterior, P: Posterior, L: Left, R: Right..... 182
- Figure 7.4.** Localisation of the documented “stress” activation region: (a) the *left PFC* (Pruessner et al., 2008), (b) the *left premotor area* (Pruessner et al., 2008). *x,y,z* values represent MNI coordinates and D: Dorsal, V: Ventral, A: Anterior, P: Posterior, L: Left, R: Right. 183
- Figure 7.5.** Localisation of the reported “stress” deactivation region: the superior temporal gyrus (Taylor et al., 1997; Pruessner et al., 2008). *x,y,z* values represent MNI coordinates and D: Dorsal, V: Ventral, A: Anterior, P: Posterior, L: Left, R: Right. 183
- Figure 7.6.** The proposed measurement architecture to monitor cortical thermal signatures. A low-cost thermal camera (directly attached to a smartphone) placed on a desk is to monitor the scalp from the front view. Another low-cost thermal camera (connected to a smartphone via a cable) is installed above the head of a person to monitor it from the top view. 185
- Figure 7.7.** Flow chart of the study protocol: stress-induction task using mathematical serial subtraction (*Counterbalanced in Latin squared design)..... 186
- Figure 7.8.** Examples of thermal images taken from (a) the frontal view thermal camera and (b) the top view thermal camera. The bounding boxes are placed on main ROIs: (a) near the left inferior frontal gyrus and (b) near the left, right PFC, left premotor, occipital lobe. 188
- Figure 7.9.** The scalp thermal imaging samples of a bald participant taken during (a) the experimental condition (i.e. the arithmetic subtraction and social evaluative pressures discussed in Section 3.3) and (b) the controlled condition (i.e. the arithmetic counting). Two circles were drawn to highlight temperature changes from the two regions (around left PFC and parietal/occipital lobes) in (a)..... 189
- Figure 7.10.** The scalp thermal imaging samples of a participant with a short hair taken during (a) the experimental condition (i.e. the arithmetic subtraction with social evaluative threat) and (b) the controlled condition (i.e. the arithmetic counting). 189
- Figure 7.11.** Significance test results visualised on each thermal image from (a) the frontal view thermal camera and (b) the top view thermal camera. The bounding boxes are placed on main ROIs: (a) near the left inferior frontal gyrus and (b) near the left, right PFC, left premotor, occipital lobe. 191
- Figure 7.12.** Box plots of 95% confidence intervals in selected (examples) metrics values across each session (control vs. experimental). Left: Metric N2 (the difference of the maximum data) from the left inferior frontal gyrus, Middle: Metric N2 from the occipital gyrus,

	Right: Metric N3 (the difference of the skewness data) from the left premotor area.	191
Figure 7.13.	The proposed measurement architecture to monitor cortical thermal signatures. To limit each participant's head motion, (a) a chin rest is installed on (b) a desk where a laptop is placed. (c) Two low-cost thermal cameras are installed above one's head to monitor (than the setup in Figure 7.6) each part of the scalp over left and right hemispheres of the brain more closely.	193
Figure 7.14.	Examples of thermal images capturing the scalp over the left and right hemispheres. The red circle with two lines inside the thermal images segments the scalp region into 6 subregions following Lombard's earlier approach (1878). More details are described in Section 7.3.2.	193
Figure 7.15.	The proposed method for extracting time-varying cortical signatures from the scalp: Step I) motion correction by selecting the scalp as a main ROI and using a ROI tracking algorithm, Step II) registration of each hemisphere by choosing three points to draw a circle (using the Hough transform), Step III) segmentation into three areas (A: anterior, M: middle, P: posterior) following Lombard (1878), Step IV) extracted cortical thermal signatures from each segmented region using basic statistical functions (maximum, average, skewness).	194
Figure 7.16.	Examples of a maximum-based time series from each ROI. Top: from a baseline, Bottom: from a controlled or experimental session.	195
Figure 7.17.	Pre-processing (interpolation & resampling and bandpass filtering) for building measures to quantify variability of the skewness timeseries from a ROI. The first measure is the number of zero-crossing points, and the second is the number of local maxima using a Matlab function <i>findpeaks</i> with a condition of 10% min peak prominence.	197
Figure 7.18.	Timeline of the London Arithmetic Stress Task (LAST): (a) the LAST has two runs and each run consists of three conditions, rest (baseline), control, experimental over the course of the nine minutes as in Pruessner et al. (2008). While the beginning condition (rest, stress or control) is counterbalanced between participants, the sequence of the three conditions is kept constant between participants: control was following rest and preceding stress. (b) A training session is included in the entire LAST protocol as in Pruessner et al. (2008).	199
Figure 7.19.	Experimental setup: with two top-view thermal cameras, an additional third camera was installed in front of a person. A finger PPG, two bracelet-based GSR/PPG sensors (worn on left, right wrists) and a breathing belt were used to build a rich physiological dataset.	201
Figure 7.20.	Boxplots (95% confidence interval) of self-reported stress scores from the 10 participants across each section (rest1, control1, experimental1 from the first run and rest2, control2, experimental2 from the second run).	202

Figure 7.21. Results from significance tests of the task type over the two runs on the metric V1 (based on zero-crossing points of the spatial skew time series) are visualised on each thermal image of the scalp over the left and right hemispheres. The right anterior and middle areas (above the right PFC and superior temporal gyrus) had significance.	204
Figure 7.22. Bar plots of the metric V1 values from the right anterior and middle areas across each run and each task condition.	204
Figure 7.23. Results from significance tests of the task type over the two runs on the metric V2 (based upon local maxima of the spatial skew time series) are visualised on each thermal image of the scalp over the left and right hemispheres. The right anterior (above the right PFC) had significance and the right posterior area (above the right parietal/occipital lobes) had an approaching significance.	205
Figure 7.24. Bar plots of the metric V2 values from the right anterior and posterior areas across each run and each task condition.	205
Figure 7.25. Results from significance tests of the task type only from the first run on basic metrics N4, N5 (Directional change in mean/maximum temperature timeseries -compared to baseline) are visualised on each thermal image of the scalp over the left and right hemispheres. The left bar plot shows the N4 values from the left anterior area across each condition. The right bar plot shows the N5 values from the right posterior area across each condition.	206
Figure 7.26. Visualisation of a <i>virtual</i> (mock-up) cortical mapping on a thermal image of the scalp to help a better registration.	209
Figure 8.1. Process for the respiration variability spectrogram collection through thermal imaging.	218
Figure 8.2. Examples of respiration spectrograms clustered according to participants' self-reported stress scores: (a) No-stress, (b) Low-level stress, (c) High-level stress.	220
Figure 8.3. Proposed data augmentation technique using a uni-directional sliding cropper with a 120x120 square window. The colour scale indicates the frequency magnitude (i.e. PSD).	221
Figure 8.4. The mainly used CNN architecture consisting of two convolutional layers, two pooling layers and one fully connected layer.	222
Figure 8.5. Flow chart of the proposed mental workload/stress induction experimental protocol: the first session is Stroop Colour Word task and the second session is the arithmetic solving task. *Tasks in each session were counterbalanced in Latin squared design.	223
Figure 8.6. k-fold leave-one-subject-out (LOSO) cross validation method: (a) dataset collected from k participants, (b) splitting the dataset into training and testing sets for each fold.	225
Figure 8.7. Boxplots (95% confidence interval) of self-reported stress scores from the 8 participants across each task (StroopEasy, MathEasy, StroopHard, MathHard).	226
Figure 8.8. Normalised subjective VAS scores: (a) distribution of the self-report per participant (95% confidence interval), (b) clustered results using k-means. Colours and shapes of each point indicate the task type and difficulty levels. (Easy: Stroop 1(S1) and Math 1(M1); Difficult: Stroop 2(S2), Math 2(M2)).	226

Figure 8.9. Test results from the three-class classification task (None, Low, High-level stress) with the CNN and the single hidden layer neural network NN1: (a) accumulated confusion matrix from 8 LOSO folds testing sets (each block consists of the sum and average over the LOSO cross validation results, the colour represents the number of each prediction (i.e., black: highest, white: lowest)), (b) accumulated accuracy from 8 LOSO fold testing sets.	228
Figure 8.10. Summary of mean classification accuracy results across 8 LOSO fold testing sets for three-class classification (i.e., none, low, high-level stress). CNN: 5 layers CNN using low resolution RVS images (28x28), NN1: single hidden layer using low resolution RVS images (28x28), NN2: single hidden layer using medium resolution RVS images (60x60), NN3: single hidden layer using high resolution RVS images (120x120).	228
Figure 8.11. Results from the binary classification task (No-stress and stress) with the CNN and the single hidden layer neural network NN1: (a) accumulated confusion matrix from 8 LOSO fold testing sets (each block consists of the sum and average over the LOSO cross validation results, the colour represents the number of each prediction (i.e., black: highest, white: lowest)), (b) accumulated accuracy from 8 LOSO fold testing sets.	229
Figure 8.12. Summary of mean classification accuracy results across 8 LOSO fold testing sets for binary classification (i.e., no-stress, stress). CNN: 5 layers CNN using low resolution RVS images (28x28), NN1: single hidden layer using low resolution RVS images (28x28), NN2: single hidden layer using medium resolution RVS images (60x60), NN3: single hidden layer using high resolution RVS images (120x120).	229
Figure 8.13. A generalised representation model for different types of physiological signals: a spectrogram generation network is to help generate variability spectrograms of each physiological pattern (e.g. blood volume pulse, etc). The output of this network can be connected to an artificial neural network for automatic feature learning.	232
Figure 8.14. Examples of pairs of the respiration variability spectrogram and blood volume pulse variability spectrogram from a rest period and a stress induction task period (the Math hard).	232
Figure 9.1. Thermal and Smartphone RGB cameras-based physiological measurement: (a) a smartphone with an add-on thermal camera, (b) flash LED emission for PPG measurement, (c) developed software to collect nasal thermal signature from the nose and BVP from the finger.	239
Figure 9.2. Heat produced by the rear flash LED of a smartphone (LG Nexus 5) measured by a thermal camera (FLIR One): (a) before turning on the LED (36.3°C), (b) after 10-15 seconds (43°C), and (c) after 25-30 seconds (53.7°C).	240
Figure 9.3. Low-level features and feature learning: we take the feature-scaled thermal variability sequence proposed in Chapter 6 as low-level features which are fed to an automatic feature learning model using a single-hidden layer neural network.	241

Figure 9.4. Experimental setup and self-report question: (a) during each stress-induction task session, (b) 20 second physiology measurement after sessions.....	242
Figure 9.5. (Boxplot) Inter-subject variability shown from the original self-reported stress scores of the 17 participants (box plot, 95% confidence interval) across each section (Rest ₁ , Stroop Easy, Stroop Hard, Rest ₂ , Math Easy, Math Hard).	246
Figure 9.6. Overall self-reported stress score distributions (from 17 participants over the sessions including the resting periods) - (a) original scores, (b) normalised stress scores (normalisation of scores from each participant) clustered into No-stress and Stress groups along with outputs of K-means.	246
Figure 9.7. Box plots of 95% confidence intervals in values of each pre-crafted metric across (a) each session (R ₁ : Rest 1, Se: Stroop easy, Sh: Stroop hard, R ₂ : Rest 2, Me: Math easy, Mh: Math hard) and (b) label produced by our labelling technique. The three features (having best correlations with self-reports) are HRV F5:RMSSD, root mean square of the successive differences of P-P intervals, Thermal F1:TD, temperature difference between from the start and the end (a red line is drawn to show negative or positive thermal direction), F2:SDTV, standard deviation of the successive differences of thermal variability sequence.	248
Figure 9.8. Summary of accumulated (from 17 fold testing sets - LOSO) confusion matrices and accuracies (a) for the three classifiers NN1, NN2 and kNN along with each set of modalities (Multimodal: HRV+Thermal, Unimodal: Thermal, HRV), (b) for the temperature-based unimodal approach built without the Rest 1 data. Each number in the confusion matrices refers to the number of instances.	250
Figure 9.9. Summary of mean inference accuracy results (testing) across 17 folds for the three classifiers NN1, NN2 and kNN along with each set of modalities (Multimodal: HRV+Thermal, Unimodal: Thermal, HRV).....	250
Figure 9.10. Summary of (a) inference accuracy along with (b) different labelling approaches (L1: K-means with k=3 and combining moderate and high stress scores, L2: K-means with k=3 and combining no-stress and moderate level stress scores, L3: K-means with k=2, L4: original scores divided by the border between no-stress and moderate levels).	252
Figure 10.1. A summary diagram of key contributions of this thesis. Red bold lines (border, arrows) indicate what we have explored and improved with respect to state-of-the-arts. Yellow boxes with red borders highlight new perspectives of our novel contributions to this research area, filling the gaps on the diagram summarising the literature in Figure 2.2.....	261
Figure 10.2. Design concept for mobile thermal imaging (hardware-aided design): a) a hardware frame attached to a chair for monitoring a person's upper body and face as in Chapter 4, b) the frame attached to a tripod for monitoring the head in Chapter 7, c) a head-set	

interface attached to a cap for the camera closely viewing a person's face in Chapter 5.....	263
Figure 10.3. Mobile thermal imaging-based in-the-wild stress intervention and management for a human worker in a factory (the EU H2020 Human research project).	264
Figure 10.4. Design concept for mobile thermal imaging (without the aid of hardware): a) a handheld concept in Chapter 9, b) the camera can be also attached to, or near, a monitor as in Chapters 6, 8, alternatively, it can be simply attached to c) a lamp, even d) a window.....	265
Figure 10.5. Mobile thermal imaging-powered anytime, anywhere with breathing monitoring (which is based upon our proposed robust respiration tracking algorithm in Chapter 5).	265
Figure A.1. Superficial muscle dissection of the neck and back.	295
Figure A.2. The task-flow of the study and an example of thermal images collected during the experiment.	298
Figure A.3. (a) thermograms collected during the push-up sessions (Participant 2), (b) ROI: thermal patterns on the upper trapezius muscle in a narrow range of temperature – just before ‘push’ (left), ‘up’ (right).	300
Figure A.4. (a) maximum temperature, (b) integral ratio based on thresholding, (c) skewness, (d) entropy on (e) a specific ROI (upper trapezius muscle) (Participant 2).	301
Figure A.5. Maximum temperature sequences on the upper trapezius muscle during the rest periods after the completion of each set of push ups (Participant 2).	303
Figure A.6. Upper trapezius muscle activations during the rest period (participant 2).	303
Figure B.1. (a) Blood vessel on the neck, (b) results: cardiac pulse rate measured by the proposed method (Grey-empty circles) and the rate measured by PPG sensor (Pink-filled circles).	304
Figure B.2. (a) Scatter plot and (b) the CAND-based accuracy histogram of cardiac pulse rates extracted through mobile thermal imaging in comparison with reference data (Infiniti finger BVP sensor).	305
Figure C.1. Extraction of eye blinks: (a) the eye-ROI on thermal image sequences during the eye-blink moment, (b) skewness along with time.	306
Figure D.1. Quantisation graphs: (a) $q=1.00 \times 2$, (b) $q=1.65 \times 2$, (c) $q=2.65 \times 2$	308
Figure D.2. A concept of Non-linear Quantisation	308
Figure D.3. Quantisation results of an image (a clenched hand and blood vessel, target area: blood vessel): (a) $q=1.00 \times 2$, (b) $q=1.30 \times 2$, (c) $q=1.45 \times 2$, (d) $q=1.65 \times 2$	309
Figure D.4. Results from the mode-seeking with the proposed non-linear quantisation	309
Figure E.1. Example materials collected in the wild: (a) materials generally found in indoor places (less influenced by environmental thermal dynamics),	311
Figure E.2. Levels of thermal dynamics of the total 14860 thermal images across the 15 indoor materials (a-o). The thermal dynamics were	

estimated by using $(T_2 - T_1)$ in Equation (4.5). [A] indicates the range of those values.....	311
Figure F.1. Overall procedure of BVP and P-P interval estimation from a person's finger through the smartphone-imaging PPG. See text for details.	313
Figure F.2. Signal extraction quality comparison of our spatial entropy-based method (Equation(2)) with the mean intensity approach (Jonathan & Leahy, 2010; McManus et al., 2013; Chan et al., 2016) by using pSQI: (a) box plot, (b) histogram.....	314

List of Tables

Table 2.1. Specification and price of commercial thermal cameras: +Low-cost, Mobile Thermal Cameras (< 34x14x68mm ³ Dimension / a mobile phone is needed for data collection), *High-end thermal imaging systems (> 90x46x56 mm ³ dimension / an additional powerful desktop/laptop is needed). ** The manufacturer corrected the thermal sensitivity values on datasheets in 2018 (from 0.15°C). In this thesis, FLIR One 2G and FLIR One Pro were used as a main channel for mobile thermal imaging.	51
Table 2.2. A summary of thermography-based affective computing studies: experimental protocols (in particular, constraints required), explored ROIs and metrics.	65
Table 2.3. Specification requirements for thermography-based physiological measurements and affective computing. The majority of existing works indeed require lower spatial and temporal resolutions than actual resolutions thermal imaging devices have.	68
Table 3.1. Proposed three stage approach based upon the research question tree in Figure 1.2.	80
Table 3.2. An overview of an incrementally built, labelled datasets from conducted data collection studies in constrained systematic and unconstrained settings	92
Table 4.1. Tracking results: the number of successfully tracked frames – Mode seeking (Cheng, 1995) vs. Median Flow (Kalal <i>et al.</i> , 2010) vs. Median Flow with the proposed Optimal Quantisation (bold: perfect).	108
Table 4.2. CNN architecture used in this study which is based on Jaderberg <i>et al.</i> (2015).	112
Table 6.1. A novel set of metrics to quantify Thermal Variability	165
Table 6.2. Significance test to assess the effect of a session type on each metrics derived from the proposed filtered sequence using a one-way repeated measures ANOVA	174
Table 7.1. Metrics for thermal directional changes based on four basic statistical measures (Average, Max, Skewness, Entropy) to represent thermal signatures from a ROI.	187
Table 7.2. A set of basic metrics to quantify cortical thermal signatures.	196
Table 7.3. A set of metrics to quantify variability of cortical thermal signatures.	198
Table 7.4. Hair types of participants for Study II.	200
Table 9.1. Pearson correlation coefficients across self-reports, PPG derived HRV and thermal high-level features. S1=Normalised self-reported scores, S2=original self-reported scores.	247

Table A.1. Comparison between differential subjective ratings based on the VAS scale and difference of maximum temperature on the upper trapezius muscle during push-ups and rest periods.....	302
Table D.1. Tracking results (based on Mean shift algorithm (Cheng, 1995)): number of successfully tracked frames without vs. with nonlinear quantisation ($q=1.45 \times 2$, 1.65×2 , 2.65×2)	310
Table G.1. Summary of feature candidates for thermal analysis.	315

Chapter 1

Introduction

“Emotional activity causes a rise of temperature (in the head); this rise is, moreover, more rapid and of greater degree than that seen in intellectual work.”

J.S Lombard (1878)

1.1. Background

With sensing technology becoming pervasive in our everyday life, the ability to monitor human psychophysiological states has become important in human computer interaction. Amongst such states, high level mental stress or mental workload is a common problem affecting mental, physical health and life in our modern society (Nash & Theborge, 2006; McEwen, 2007; Arnsten, 2009). Studies show that mental stress could be automatically assessed through the use of these physiological sensing technologies, in turn opening new potential ways for stress management support strategies (Healey & Picard, 2005; Hosseini & Khalilzadeh, 2010; Hernandez *et al.*, 2011; Hong *et al.*, 2012; Sano & Picard, 2013; Al-Shargie *et al.*, 2016; Yu *et al.*, 2018). However, to date, such available technologies are relatively heavy, fragile, or restrict mobility, movement and measurement environments, limiting their use. This thesis focuses on building new approaches to more reliable physiological and stress measurements using low-cost and mobile sensing technology, supporting unconstrained and potentially a variety of everyday situations.

Mental stress is a complex phenomenon given its definition, “*the nonspecific response of the body to any demand*” noted by Hans Selye who first introduced the term stress in 1926 (see his exemplary guidance to stress research, Selye, 1980). As stress influences a person’s neural and physiological processes, various approaches have been proposed to extract signatures of the human stress response. Sensing channels for cardiovascular (e.g. heart rate, heart rate variability, blood pressure, vasoconstriction/dilation-induced temperature), respiratory (e.g. breathing rate,

tidal volume), perspiratory (e.g. skin conductance level and response rate), muscular (e.g. muscular electric wave, muscle activation-induced temperature) and neural activities (e.g. brain electrical wave, cerebral blood flow activation) have been employed as physiological measures of stress (Healey & Picard, 2005; Hosseini & Khalilzadeh, 2010; Hernandez *et al.*, 2011; Hong *et al.*, 2012; Engert *et al.*, 2014). To ensure high performance, existing works require direct or close contact to the body, often using cumbersome attachment mechanisms (e.g. chest belt, EEG cap) and often highly constrained settings, limiting its deployment in real world applications.

Interestingly, like systems for capturing behavioural data in a contactless manner, such as body motion (Kleinsmith & Bianchi-Berthouze, 2013) and voice (Lane *et al.*, 2015), researchers have shown that affect-related physiological signatures can also be captured via non-contact sensing devices, such as thermal cameras (e.g. respiration monitoring in Pereira *et al.*, 2015) and RGB cameras (as remote photoplethysmography to measure blood volume pulse in Verkruijsse *et al.*, 2008). Whilst RGB camera-based methods suffer from illumination and privacy issues, thermography is much less affected by those constraints (Lloyd, 2013). In addition, studies have shown that different types of physiological activities can be read through the use of a thermal imaging channel (Garbey *et al.*, 2007, Pavlidis *et al.*, 2012, Pereira *et al.*, 2015).

Thermography, also called thermal imaging, which interprets electromagnetic radiation emissions into temperatures and visualises them on a two-dimensional image, is a key non-contact and non-invasive method that helps to capture heat distributions over human cutaneous skin regions. With advancements in commercial thermal imaging technologies, thermography has supported studies to explore the relation of a person's skin temperature with his/her affective or mental states (an example of facial thermal images of the author is shown in Figure 1.1a, which was taken by a high-precision thermal camera in Figure 1.1b). A variety of studies have investigated patterns of temperature variations (e.g. thermal direction) associated with psychological states. For example, Engert *et al.* (2014) explored thermal responses of a person's facial regions to mental stressors, finding a significant drop in nasal temperature due to stress-induced vasoconstriction of the area. Computational physiologists have also explored capabilities of thermal imaging in extracting other types of physiological signals (e.g. respiratory rate in Pereira *et al.*, 2015 and cardiac pulse rate in Garbey *et al.*, 2007).

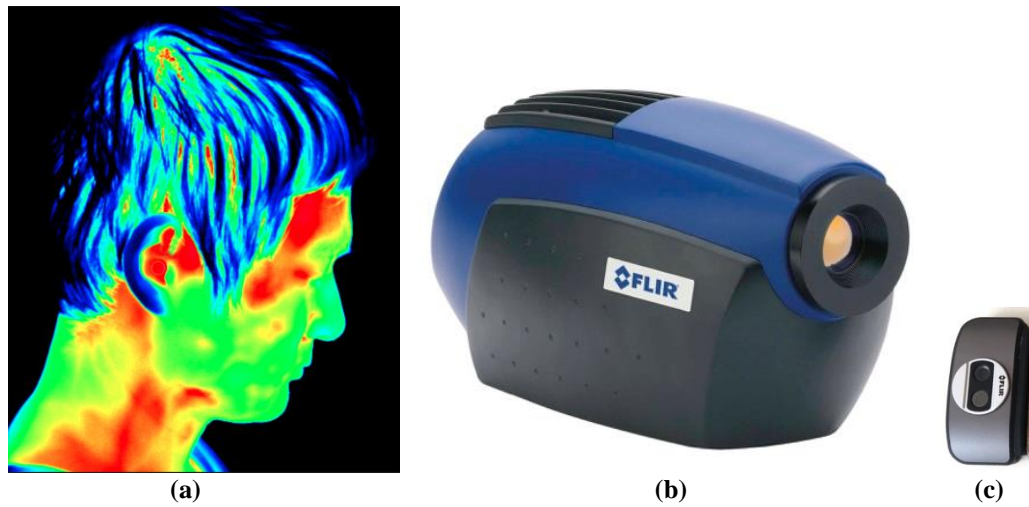


Figure 1.1. (a) an example of a thermal image of the author taken by (b) a high-precision, high-cost, heavyweight thermal camera (FLIR SC5000mb) – dimension: 320x141x159 mm³; weight: 3,800.0g (this camera also needs to be used with a powerful desktop/laptop for data collection). And (c) a low-precision, low-cost, lightweight thermal camera (FLIR One 2G) – dimension: 34x68x14 mm³; weight: 36.5g

Although earlier initial investigations have shown promising capabilities of thermal imaging in assessing affective states and measuring a person's physiological signals, the findings in the literature are mostly limited to discussions of binary directions of thermal variations (thermal directionality) on facial and palm regions in association with certain affective states (Genno *et al.*, 1997; Pavlidis *et al.*, 2002, 2012; Gane *et al.*, 2011; Engert *et al.*, 2014). Its simplicity makes it susceptible to noise due to many external (e.g. environmental temperatures) and internal (e.g. body conditions) factors, being likely to lose important information as often reported in studies using other physiological channels (Billman, 2013).

In addition, performance of physiological measurements using thermal imaging has been still very limited as they require highly constrained settings in terms of room temperatures and a person's mobility. This has been one of factors limiting the capability of physiological thermal signatures in assessing a person's affective states. In particular, despite other types of thermal signatures discovered such as respiratory thermal signatures, only the simple directional information has been considered in assessing mental stress, often resulting in inconsistent findings. For example, while Engert *et al.* (2014) reported a significant temperature drop of the chin in response to mental stress, Veltman & Vos (2005) reported there was no significant change

in temperature of the same area. Furthermore, all of the earlier studies have employed heavyweight and expensive thermographic systems (e.g. Figure 1.1b – cost: over £10,000 in 2017, weight: 3,800g, dimension: 320x141x159mm³), which can only be set up in certain positions and require to be used with additional powerful computing machines for data collection, hence limit their use to highly constrained laboratory settings, in turn making themselves less likely to be used in mobile, everyday HCI applications.

Now, we are standing at an exciting moment when advanced thermography technology has emerged, producing a new category of thermographic systems: *mobile, low-cost thermal imaging system* (e.g. Figure 1.1c – cost: lower than £200 in 2017, weight: 36.5g, dimension: 34x68x14 mm³). Despite the relatively low quality of their thermal imaging outputs, this technology could help bridge the gap between the preliminary findings from highly constrained laboratory environments and real-world applications in the wild. Indeed, its portability (e.g. small size and low computational resource requirement) allows it to not only be easily attached to mobile phones but also be integrated into our clothes and accessories. Inspired by the initial explorations in the body of earlier work of thermal imaging as a measure of physiological and affective states, this research aims to explore how to bring *Thermography* into unconstrained, real world human computer interaction as a multimodal physiology and mental stress sensing channel.

1.2. Research Questions, Scopes and Contributions

To achieve the aim addressed above, we discuss our research questions, the scope and contributions of this thesis in this section. Figure 1.2 shows our research question tree starting with a key question, “*Can mobile, low-cost thermography be used as multiple physiological measures for automatic recognition of a person’s mental stress in HCI settings?*”. We discuss each detail below.

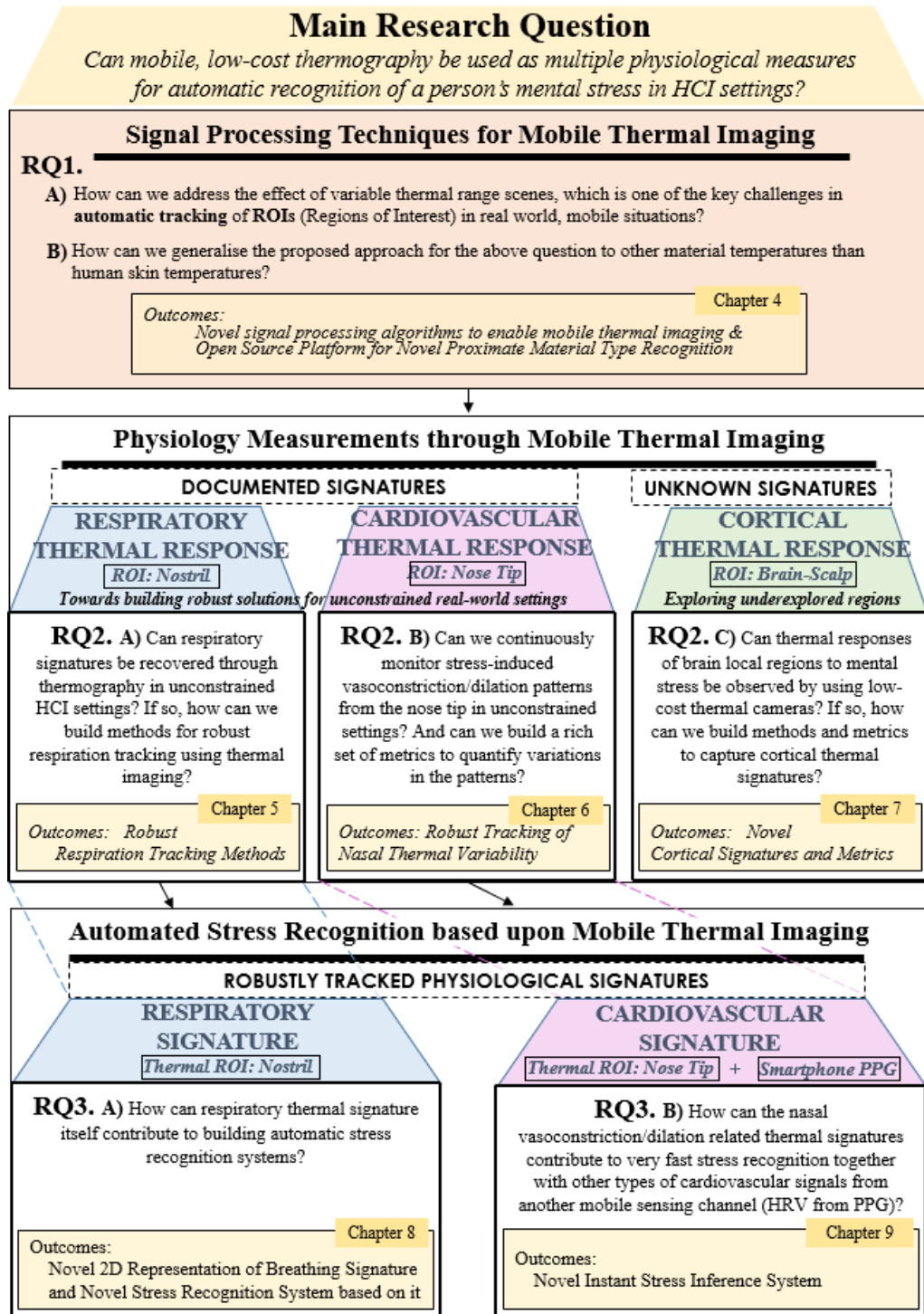


Figure 1.2. Research Question (RQ) Tree: the main research question leads to detailed sub-questions along with three themes: i) signal processing techniques for mobile thermal imaging, ii) physiology measurements through mobile thermal imaging, iii) automated stress recognition based upon mobile thermal imaging

Improving Signal Processing Techniques for Mobile Thermal Imaging

RQ1. A) *How can we address the effect of variable thermal range scenes, which is one of the key challenges in automatic tracking of ROIs (Regions of Interest) in real world, mobile situations?*

The ultimate goal of this work is to formulate methods which can bring mobile thermal imaging into a variety of health-related real-life applications. Until recently, thermographic systems have been regarded as infeasible and impractical solutions for physiological and psychological measurements because they were very costly and heavy as highlighted by Gastel *et al.* (2016). The emergence of recently launched mobile thermographic systems, which are low-cost and small-sized (e.g. Figure 1.1c), opens the possibility of their use in unconstrained, ubiquitous mobile settings. However, to have mobile ubiquitous systems that operate in the general environment and not merely controlled indoor-laboratory settings we ought to confront the key challenge: highly-variable environmental temperature which negatively affects morphological, graphical properties of thermal images in automatically tracking a Region-Of-Interest (ROI), which is a prerequisite to reliably and continuously extract physiological signatures.

To overcome the limitations, we propose new quantisation-based pre-processing techniques for enabling existing ROI tracking methods to properly perform in uncontrolled thermal range scenes. The performance of our approach is evaluated by comparison with state-of-the-art methods without the proposed quantisation in settings of incremental complexity in terms of mobility and variability of environment temperature. The main concept is proposed and verified in Chapter 4 and further investigations are made based upon this method in accordance with different ROIs in Chapters 5 and 6.

RQ1. B) *How can we generalise the proposed approach (above) to other material temperatures than human skin temperatures?*

Answering the question, we aim to demonstrate the general capability of the proposed quantisation-based pre-processing to other material temperatures in real world settings. In particular we investigate how it can contribute to building another HCI application of thermal imaging beyond physiological and affective computing applications which is the focus of this thesis. Thermal imaging has also been investigated for other HCI applications. For instance, Larson *et al.* (2011) proposed a thermal imaging-based surface gesture recognition system which uses residual heat traces on a surface. Kurz (2014) extended this concept to support touch

detection in 3D augmented space. Shirazi *et al.* (2014) have shown how to use thermal reflection characteristics of materials for gestural interaction. However, these approaches have also limited their use to constrained settings due to the use of high-cost, heavyweight thermal cameras. Given this (even in other HCI applications), the issue due to environmental temperature changes have been unaddressed up to this point. To address this, the last part of Chapter 4 targets a specific task, classification of material types collected outdoors where ambient temperature is extremely dynamic. Despite computational methods built for this not being directly used in the following chapters, the findings provide general insights of the significant role of the quantisation process.

The main outcome from the question RQ1(A,B) is the base for physiological computing and automatic stress recognition following from the structure of this thesis. Software libraries built for these methods are publicly available on web-based archives (e.g. GitHub) to foster the approaches in this research community.

Enabling Mobile Thermal Imaging as Multiple Physiological Measures

Amongst three questions below to address the challenge, the first two questions are about how to achieve robustness in monitoring documented physiological responses, i) respiratory and ii) cardiovascular thermal responses in unconstrained settings. This selection of the two responses is particularly important due to both being the most rigorously investigated in the literature. At the same time, tracking them reliably in unconstrained settings have been underexplored, which will be tackled within this thesis. The remaining question among the three is about an investigation of a new cortical thermal response from underexplored regions. This is to tackle another limitation of the body of earlier work whose findings have been primarily from facial regions.

RQ2. A) *Can respiratory signatures be recovered through thermography in unconstrained HCI settings? If so, how can we build methods for robust respiration tracking using thermal imaging?*

This thesis focuses on how to reliably recover respiratory signatures from the breathing induced air exchanges inside the nostril which is the key region where we can observe thermal changes along with inhalation and exhalation cycles (Pereira *et al.* 2015). By doing so, we further identify key issues in unconstrained settings: difficulty in automatically tracking the nostril due

to noise coupled by motion artefacts (mobility) and thermal dynamics (varying ambient temperature), and the degraded quality of respiration signals due to the noise together with the low resolution of mobile thermal imaging. Hence, we propose novel methods to robustly track the nostril and respiratory signature in Chapter 5. The proposed methods are thoroughly evaluated in three studies which are conducted with different levels of complexity in terms of the issues addressed. The contribution of the robustly tracked respiratory thermal signature to stress recognition is thoroughly investigated in Chapter 8 where a breathing-based stress inference system is built.

RQ2. B) *Can we continuously monitor stress-induced vasoconstriction/dilation patterns from the nose tip in unconstrained settings? And can we build a rich set of metrics to quantify variations in the patterns?*

While binary directions in thermal variances on a facial ROI (e.g. nasal) occurring in association with a person's affective state has been the main focus in this field of study, computational algorithms for supporting the analysis (e.g. automated tracking of a ROI) have been very limited in a majority of earlier works (e.g. a dot stick was used for the tracking of the ROI in Nhan & Chau, 2010). In particular, although it has been repeatedly shown that the vasoconstriction/dilation-related temperature drop from the nose tip occurs in response to mental stressors, there has been a lack of methods to continuously and reliably capture the entire pattern of cardiovascular signature rather than just a binary direction. As discussed above, the simplicity of the metric measuring the direction makes it susceptible to noise due to a variety of external and internal factors such as variable thermal range scenes and physical, psychological conditions of a person, possibly losing important information.

Inspired by a controlled study which has found the very quick response of the signature under mental load (Abdelrahman et al. 2017), in Chapter 6 we propose computational methods and metrics to automatically track the signature and quantify even subtle, but informative variations in it from where head motion and environmental temperatures are not controlled. The proposed methods are thoroughly evaluated in a controlled systemic experiment and an unconstrained stress induction study. The contribution of the robustly tracked nasal vasoconstriction/dilation related thermal signature to stress recognition is thoroughly investigated in Chapter 9 where an instant stress detection task is used.

Beyond the known physiological thermal signatures which are mostly from our facial area which is explored in the literature, this thesis investigates underexplored regions which could show thermal responses to stressors. Based on evidence from the human physiology on heat production, the scalp is chosen as a new ROI to identify its cortical thermal response. In detail, this thesis aims to answer:

RQ2. C) *Can thermal responses of brain local regions to mental stress be observed by using low-cost thermal cameras? If so, how can we build methods and metrics to capture cortical thermal signatures?*

Various local regions over the scalp can be monitored using mobile thermal imaging by taking advantage of it being lightweight and portability by contrast with dominantly used heavyweight devices. A low-cost, mobile thermal camera can be installed above a person's head easily and this recording setup does not require to obstruct one's vision in real life situations. With this, the thesis builds new methods for cortical mapping on thermal images to find strong cues related to brain activation (i.e. cerebral blood flow) in Chapter 7. For this, we search for neurological and physiological evidence of heat production of the brain in response to mental stressors and then elaborate stress-related local brain regions for the choice of ROIs.

A computational mapping procedure can be built on the automatic ROI tracking methods proposed in the earlier chapters, which provide the positional information of a ROI. However, the scope of this work does not include the design of a new tracker. We also investigate a wide variety of metrics to quantify cortical thermal signatures. The investigations and evaluations are undertaken in an unconstrained stressful situation to lead initial discussions and a highly constrained stress induction study which minimises motion and hair-induced noise which may contaminate signatures. Given the aiming at initial explorations of new cortical thermal signatures, the outcome is not further extended in the following chapters where stress recognition systems are built.

With the outcomes from the proposed physiological computing approaches, this thesis finally aims to build automated stress recognition models and systems. In particular, we investigate the contribution of each robustly tracked physiological signature (i.e. respiratory and

cardiovascular thermal signatures) to mental stress inference since their roles have been weak or underestimated. With this purpose, the next question is:

Automatically Recognising a Person's Mental Stress

RQ3. A) *How can respiratory thermal signature itself contribute to building automatic stress recognition systems?*

In the computing literature of automated stress recognition, it has been shown that the contribution of traditional gross respiration metrics (e.g. average breathing rate for a certain duration) to automatically recognising stress is weak in comparison with other physiological features (e.g. heart rate variability). This may be due to limitations of the widely used metrics simplifying the physiological information. Indeed, Grossman (1983), a physiologist, addressed earlier, “*stressful events could affect a person's respiration irregularity*”. Inspired by his note, this thesis proposes a new bi-dimensional way to represent respiration signature which can condense dynamic respiratory information. This allows us to conduct the in-depth analysis of respiration variability in relation to mental stress in Chapter 8. Furthermore, a deep learning-based framework is designed for automatic feature learning on the proposed signature. Consequently, the thesis presents a novel respiration-based unimodal stress recognition system trained on a dataset we have collected from a stress induction study.

The next and the last question is:

RQ3. B) *How can the nasal vasoconstriction/dilation related thermal signatures contribute to very fast stress recognition together with other types of cardiovascular signals from another mobile sensing channel (HRV from PPG)?*

To jump up one more step, this thesis aims to investigate instant stress recognition (i.e. inferring mental stress levels from a very short period of physiological measurements – several seconds) considering that a variety of systems must take a relatively long term window of data (several minutes to a few hours) to ensure reliable inferences (McDuff et al., 2016), which may not be feasible in real world situations. In Chapter 9, we investigate the contribution of the nasal vasoconstriction/dilation related signatures to this particular task. Here, we also take a look at other types of cardiovascular signals, photoplethysmography, which can be simultaneously

obtained from a mobile phone-based setup (which is a controller of mobile thermal imaging). The practicality and feasibility of the proposed system is tested in unconstrained stress inference tasks.

Answering the organised research questions, this thesis finally proposes *Mobile Thermography* as a new multi-modal physiology and stress sensing channel, which can be potentially used to tailor physical and mental activities to a person's affective needs in real world applications.

1.3. Thesis Structure

The rest of the thesis is organized as follows:

In Chapter 2, a literature review is conducted to identify (a) a past use of temperature measurements and available commercial thermographic systems, (b) how thermal imaging can be used to measure physiological signatures (e.g. direction of thermal variations and respiration), and earlier challenges, (c) earlier findings on affect (including stress)-related thermal signatures, and (d) explorations of human physiology and neurology on stress-related heat production to find new regions of interest over the scalp, which respond to stress.

In Chapter 3, the main research methodology is proposed together with a summary of basic tools for mobile thermal imaging, experimental protocols for incrementally building datasets, and evaluation methods.

We start looking at **Stage I** (Chapter 4). First, we propose a new thermal image pre-processing method, *Optimal Quantisation*, which can minimise negative effects of ambient temperature on automated tracking processes of ROIs on the human skin. Second, we demonstrate the general capability of the approach to other material types than the human skin to show how the method can be used in other HCI applications of thermal imaging (Chapter 4).

At **Stage II** (Chapters 5, 6 and 7), we dive into three different physiological thermal signatures. The first two chapters focus on the two documented physiological signatures in the literature: respiratory signature (Chapter 5) and nasal vasoconstriction/dilation-related cardiovascular signature (Chapter 6). In particular, we explore how to achieve reliable accuracies in computationally tracking each signature in unconstrained, mobile and ubiquitous situations

(e.g. walking outside in Chapter 5). The remaining chapter (Chapter 7) explores novel cortical thermal signatures from underexplored areas (i.e. local regions of the scalp) in relation to cortical activation and deactivation of local cortical areas associated with mental stress.

At **Stage III** (Chapters 8, 9), we explore the contribution of each physiological thermal signature to stress detection. Particularly, we focus on respiratory and cardiovascular signatures as they are able to be robustly tracked using the proposed methods in the second stage. As the computing literature has shown that their roles have been weak in this particular task and incongruent results have been often reported, we investigate how to address the issues by focusing on each signature separately rather than combining all signatures together to obtain better performance. With this focus, we propose a non-contact stress recognition model which is based upon a newly proposed two-dimensional respiration signature in relation to mental stress levels (Chapter 8), and then we explore the contribution of the nasal vasoconstriction/dilation related thermal signatures to a very fast stress recognition procedure together with other types of cardiovascular signals from another mobile sensing channel (Chapter 9).

Finally, this thesis ends with conclusions in Chapter 10. In the next chapter, we dive into research literature in the areas spanning from the use of earlier thermometry, thermal imaging, to human physiology in relation to stress-related heat production.

Chapter 2

Background

As humans are homeothermic, our internal temperature is closely linked with numerous physiological and psychological mechanisms. Given this, human thermal patterns have been explored to improve the understandings of our body for a couple of centuries. This thesis is also concerned with this, particularly, how to more reliably interpret a person's temperature into physiological signatures and help detect our psychological stress, possibly in any situations. Through this chapter, we dive into existing methods, paradigms and physiological evidence around the thermal mechanism. First, we start looking at thermal measurements spanning from earlier thermometry to modern thermal imaging, and their use. Second, we discuss recent studies exploring the use of thermography in capturing a person's physiological cues and understanding affective states through thermography. Then, we review human physiology of heat production to discuss an underexplored thermal body area and its related signatures. Finally, we state the challenges and limitations which have emerged from the literature.

2.1. An Overview: from earlier Medical Thermometry to state-of-the-art Thermal Imaging

A brief history of medical thermometry

Back to the eighteenth century when the first mercury thermometer was invented by D.G. Fahrenheit in 1714 (Fahrenheit, 1724), physiologists and philosophers had begun to explore bodily heat of a living mammal in relation to medical symptoms (Pearson, 1786; Crawford, 1788). Authors observed that certain types of diseases or inflammatory disorders could induce an increase of the temperature of the body. In the nineteenth century, the bodily temperature of

patients had been more rigorously investigated; in turn, the heat monitoring had started to be used as a diagnostic and prognostic tool for fatal diseases such as tumours and phthisis (Jackson, 1828; Ringer, 1873; Donkin, 1879). In the late nineteenth century when physiologists brought up the underexplored topic of the cognitive, sensorial, and emotional functions of the brain to the research community (James, 1884), their earlier attentions given to the body temperature had been extended to the heat from the brain. In 1877, to the best of the author's knowledge, Broca applied thermometry to the surface of the head, for the first time, to estimate the mean temperature of the brain (Anon, 1877; Haller Jr, 1985)⁶. A few years later, Lombard investigated the head temperature under *cognitive* and *emotional* conditions (Lombard, 1878a, 1878b). The author found increases of temperatures from three segmented regions of the head: anterior, middle and posterior.

Explorations on the temperature of the human body and the brain had continued throughout the first half of the twentieth century in which thermometry became more accurate and more general in use; in particular, psychiatrists and neurologists had studied the effect of affects and emotions on the bodily temperature of a human being (Ziegler & Cash, 1938) and emotional hypothermia in an animal (Grant, 1950). These earlier discoveries with the thermometry over the previous centuries paved the way for the understanding of what temperature changes of our body regions would mean; however, it was almost impractical to observe temperatures from the entire skin surficial area which may have different patterns until a new type of thermometry was invented and commercialised around the middle of the twentieth century – that is, *the infrared thermal imaging camera* (see the historical review of the thermal imaging devices; Lloyd, 2013).

Thermal imaging (thermography) Technology

Non-contact thermography is a key non-invasive method to study heat distributions on the surface of materials and organisms based on the interpretation of naturally emitted electromagnetic radiations into temperatures. Most commercial thermographic cameras sense the electromagnetic radiations (the wavelength range: between approximately 8µm and 14µm) from surfaces of objects and materials, providing thermograms of them. Thermograms can have a wide range of spatial and temporal resolutions similar to images captured by RGB-vision cameras.

⁶ As Broca's original paper was not available online in 2015-2018 when the thesis was written, we referred to two review articles published in 1877 (The Lancet) and in 1985 (The Western Journal of Medicine).

Here, the ability to read electromagnetic radiations outside the visible spectral range (so-called visible spectrum) differentiates their benefits from the visible spectrum-based imaging devices, such as RGB cameras, which are susceptible to *illumination effects*: for instance, extremes of darkness and brightness incapacitate imaging capabilities due to sensor saturation or sensitivity (Liu *et al.*, 2010; Aittala *et al.*, 2015).

In detail, human eyes and RGB cameras that mimic the eyes are only sensitive to a narrow range of electromagnetic spectrum called “visible light”. They capture objects through the radiations in the visible spectrum their surface emits or reflects when hit by a source of light (Roychoudhuri *et al.*, 2008). This means that without sources emitting visible light, such as the sun or incandescent bulbs, this type of vision system is generally unable to “see” the world. According to *Planck’s law*, every object above absolute zero emits thermal radiation (Pavlidis *et al.*, 2000; Vollmer & Klaus-Peter, 2017). Most of emitted radiations fall in the infrared spectrum range (circa between 0.9 and 14 μm) rather than the visible spectrum range (approximately between 380 and 780 nm). Thermal cameras are designed to capture infrared radiations while RGB cameras are not.

A wide range of factors influence capturing of thermal radiation patterns emitted from materials. Some examples are compound combinations of materials, surface structure, such as the level of roughness, and geometry such as existence of cavities. Thus, what thermal imaging systems measure is actually estimated temperature using some thermal parameters, such as, the emissivity which is the level of effectiveness of materials in emitting energy as thermal radiation (Vollmer & Klaus-Peter, 2017)

Since the human body is homoeothermic, and cutaneous and subcutaneous skin regions release infrared radiations, they can be remotely interpreted and transformed into thermograms (thermal images) (Lloyd, 2013). Such thermograms produced by thermal imaging systems have been used in medical applications to detect pathological symptoms, disorders and diseases of patients. Clinical studies using thermal imaging have demonstrated its high performance in detecting inflammatory arthritis, osteoarthritis, soft tissue rheumatism, and malignant diseases or tumours, and a good performance in clinical monitoring of Complex Regional Pain Syndrome (CRPS) and Raynaud’s phenomenon, whose symptoms include abnormal or asymmetric temperature distributions visually inspectable by clinicians (Wasner *et al.*, 2002; Spalding *et al.*, 2008; Ring & Ammer, 2012).

With advancements in commercial thermal imaging technologies, thermography has been adopted beyond medicine to investigate the possibility of measuring physiological processes

(Murthy & Pavlidis, 2006; Garbey *et al.*, 2007; Lewis *et al.*, 2011; Ebisch *et al.*, 2012; Gault & Farag, 2013) and affective states (Puri *et al.*, 2005; Or & Duffy, 2007; Pavlidis *et al.*, 2012; Abdelrahman *et al.*, 2017). However, such studies have a common limitation. Their findings have all relied upon the use of *high-cost*, heavyweight (several kilograms) thermographic systems (with their dependency on an additional powerful computing machine) in highly constrained settings, in turn keeping themselves from being deployed to unconstrained, mobile, real world-like settings (We discuss details of these studies, their findings and limitations in the following sections).

Very recently, a new group of *low-cost*, lightweight (<37gram) thermal cameras have been developed, which can be used as an accessory of mobile devices (e.g. smartphones, tablet PCs). They have greater potentials to power real world applications of thermal imaging (a smartphone embedding a thermal camera has been already launched⁷). For example, Naik et al. (Naik & Patel, 2017) have shown that, in controlled environments, a low-cost thermal camera has sufficient resolution and accuracy to grade the maturity of a fruit.

Lastly, Table 2.1 summarises specifications and costs of both *low-cost* and *high-cost* thermal imaging devices between 2015 and 2018. There are three key parameters in determining thermal imaging quality: thermal sensitivity, spatial resolution and temporal resolution (sampling rate). The thermal sensitivity is measured by Noise Equivalent Temperature Difference (NETD). NETD is to describe the minimal temperature difference which can be recognisable by a thermal imaging device. Hence, the smaller, the better. NETD of thermal cameras varies from 0.017°C to 0.5°C. At this very moment, the cost of the currently available sensors seems to primarily depend on the temporal resolution (here, sampling rate). For example, while both FLIR One (at £166 GBP, in May 2016) and Optris PI 200 (at £2,500 GBP, in May 2016) meets the 160x120 spatial resolution and under-0.1°C sensitivity, the former has unsteady sampling rate lower than 8.7Hz and the latter supports 128Hz. Nonetheless, apart from specifications, such low-cost cameras (+ in Table 2.1) provide valuable benefits, such as small form factors, affordability and portability (e.g. lower than 34x14x68mm³ dimension), making them more feasible to be deployed in HCI systems than heavyweight, immobile, and expensive high-end thermographic systems (* in Table 2.1).

⁷ CAT S60: <https://www.catphones.com/>

Table 2.1. Specification and price of commercial thermal cameras: +Low-cost, Mobile Thermal Cameras (< 34x14x68mm³ Dimension / a mobile phone is needed for data collection), *High-end thermal imaging systems (> 90x46x56 mm³ dimension / an additional powerful desktop/laptop is needed). ** The manufacturer corrected the thermal sensitivity values on datasheets in 2018 (from 0.15°C). In this thesis, FLIR One 2G and FLIR One Pro were used as a main channel for mobile thermal imaging.

Product	Spatial Resolution	Sampling Rate	Thermal Sensitivity (NETD)	Price (in July 2017)
+FLIR One 2G (Figure 1.1c)	160x120	<8.7Hz (unsteady)	<0.10°C	Not applicable (in July 2017) £166.00 (about \$239 in May 2016)
+FLIR One 3G	80x60	<8.7Hz (unsteady)	0.10°C**	\$199.99
+FLIR One Pro	160x120	<8.7Hz (unsteady)	0.07°C**	\$399.99
+Seek Thermal Compact	206x156	<9Hz (unsteady)	0.50°C	\$249.00
+Seek Thermal Compact PRO	320x240	>15Hz (unsteady)	Unknown	\$499.00
+Therm-App Hz	384x288	25Hz	<0.07°C	\$999.00
*Optris PI 200	160x120	128Hz	<0.10°C	£2,500.00 (about \$3,570 in May 2016)
*FLIR A35	320x256	60Hz	<0.05°C	\$5,900.00
*OPTRIS PI-640	640x480	32Hz	0.075°C	\$11,310.04
*FLIR A325sc	320x240	60Hz	<0.05°C	\$12,322.46
*FLIR SC5000mb (Figure 1.1b)	640x512	100Hz	0.017°C	Over \$13,000.00

2.2. Thermography as a Measure of Physiological Cues

In this section, we review existing works about the use of thermal imaging in measuring a person's physiological cues. We are particularly interested in what types of physiological thermal signatures have been explored, which methods and metrics have been developed and what has

been required in the prior art for extracting and quantifying such signatures through thermal imaging.

The Autonomic Nervous System (ANS) controls physiological phenomena, such as, cardiovascular, respiratory, perspiratory activities as well as muscular activity which is also linked to behavioural patterns. Such physiological cues have been employed as measures of a person's affective states, in particular, mental stress (Healey & Picard, 2005; Hosseini & Khalilzadeh, 2010; Hernandez *et al.*, 2011; Hong *et al.*, 2012; Engert *et al.*, 2014). Recent studies using thermography have shown possibilities of thermal imaging in monitoring human physiological signatures from facial areas. The most widely explored physiological thermal cue is the directional change in temperature which is triggered mainly by vasoconstriction and vasodilation related cardiovascular activity (Genno *et al.*, 1997; Or & Duffy, 2007; Engert *et al.*, 2014; Di Giacinto *et al.*, 2014; Salazar-López *et al.*, 2015).

Vasoconstriction/dilation related cardiovascular thermal signature

In particular, the nose tip area from which a variety of studies have explored temperature patterns (Genno *et al.*, 1997; Veltman & Vos, 2005; Or & Duffy, 2007; Di Giacinto *et al.*, 2014; Engert *et al.*, 2014; Abdelrahman *et al.*, 2017) is one of the main facial areas where we could observe vasoconstriction and dilation induced thermal directional changes. Vasoconstriction over the area induces a decrease in temperature. This is the narrowing of blood vessels, causing blood flow to decrease and in turn reducing loss of body heat. Hence, the surface area becomes colder. This can occur under a cold ambient temperature condition (Pergola *et al.*, 1993) as well as a specific affective (e.g. mentally stressed) condition (Elam & Wallin, 1987). Vasodilation occurs in an opposite way. Given this, most researchers in this domain have focused on capturing a directional change (i.e. temperature drop, rise) in temperature⁸ of the nose tip as a metric.

Capturing has been mainly done in three ways: i) subtracting temperature from the final moment of a task from the initial moment of the task (Or & Duffy, 2007), ii) comparing temperature from a task with that of another task (Di Giacinto *et al.*, 2014; Salazar-López *et al.*, 2015; Abdelrahman *et al.*, 2017), and iii) extracting a linear slope from an entire temperature sequence from a task (Engert *et al.*, 2014). As this procedure is of relative ease, a support of

⁸ To represent temperature of a Region Of Interest (ROI), simply averaging temperatures in the ROI (here, nose tip) has been dominantly used (Veltman *et al.*, 2005; Or & Duffy, 2007; Di Giacinto *et al.*, 2014; Engert *et al.*, 2014; Abdelrahman *et al.*, 2017). Note that thermographic systems provide two dimensional temperature matrices, so multiple temperatures can be observed from the nose tip.

complex computational methods for automatically continuously extracting thermal variations has not been necessarily required. For example, Engert *et al.* (2014) visually inspected and rejected thermal images where there is a movement of a participant. To reduce this effort, many authors have requested their participants to stay in an exact same position (Salazar-López *et al.*, 2015). A chinrest has also been often used to make this process easier (Veltman & Vos, 2005). Although a few studies have used computational tracking methods, the purpose has been to correct minor motion artefacts as participants have been generally asked to take a look to the front facing a thermal camera (Abdelrahman *et al.*, 2017). To ensure the minimisation of environmental temperature effects on this measure, room temperatures have been generally controlled (Genno *et al.*, 1997; Or & Duffy, 2007; Abdelrahman *et al.*, 2017). Such constraints have been generally introduced to studies in the literature reviewed in this section as well as Section 2.3 (later, we provide tables to summarise the literature in Section 2.3.3).

Perspiratory thermal signature

Depending on the selection of a proper ROI, different types of physiological thermal signatures can be monitored. For instance, skin regions containing many sweat glands, such as the palm, the maxillary and the perinasal regions can have perspiratory responses (i.e. sweat gland activations). The sweat gland responses result in either an increase or a decrease in temperature of those areas (Shastri *et al.*, 2009; Pavlidis *et al.*, 2012; Krzywicki *et al.*, 2014). It is noteworthy that Pavlidis *et al.* (2012) identified strong correlations between galvanic skin response (GSR) signals and thermal changes both on the perinasal area and on the finger region ($r=0.943$, $r=0.968$, respectively) during psychological stress inducing tasks. More recently, Krzywicki *et al.* (2014) proposed a method to find active pores on the cutaneous skin and count them using thermal imaging. As shown in Figure 2.1, extremely high-resolution thermal cameras (640x512, 1280x1024) with focal lens enabled the authors to observe perspiratory activations of minute pores (i.e. sweat glands). In addition, Shastri *et al.* (2009) highlighted noise from (automatic) tracking and segmentation imperfections, leading to low accuracies in extracting such physiological signatures.

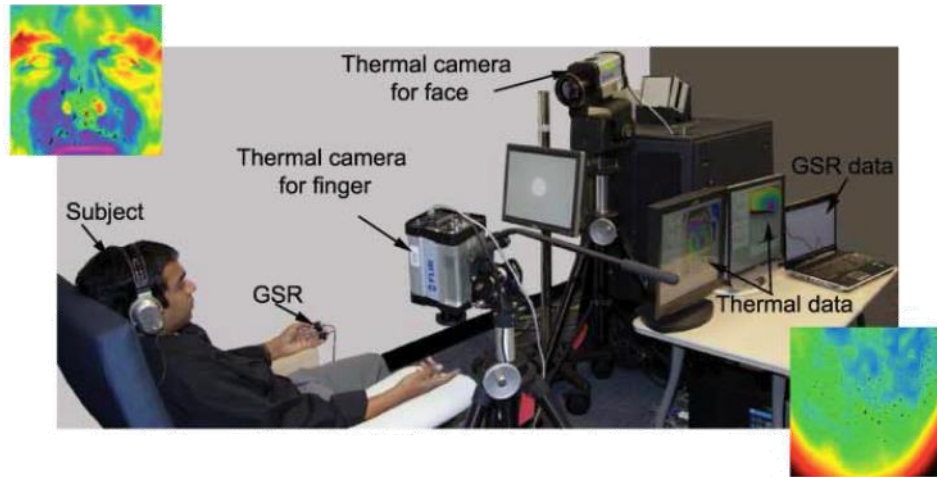


Figure 2.1. A widely used experimental setup in the literature (adapted from Pavlidis *et al.* 2012).

Cardiac pulse related cardiovascular thermal signature

The superficial blood vessels play an important role as an indicator of human health, providing a measure of temperature as well as cardiac pulse-related information, due to vessels being directly influenced by cardiac output and arterial pressure. Like the computation for extracting sweat glands responses, the cardiac pulse related information can be extracted through the use of high-precision and expensive thermographic systems as shown in Figure 2.1. Garbey *et al.* (2007) hypothesised that the cardiac pulse could be recovered from the dynamic thermal signal emitted from major superficial blood vessels since the temperature on the vessel is modulated by the blood flow. The authors proposed an algorithmic approach mainly based on a conditional density propagation tracker (Isard & Blake, 1998) to track a selected artery vessel on thermal images. They estimated cardiac pulse rates by using the fast Fourier transform (FFT).

However, their experiment comparing the extracted pulse rates with reference data showed very low accuracy; for example, although the range of participants' pulse rates (ground-truth) was very narrow (55.6bpm – 86.1bpm), the maximum error was 23.7bpm⁹. The authors also documented that the performance relied upon thermal imprint quality, which is entirely dependent on the ROI tracking performance. Furthermore, by contrast with other types of cardiac

⁹ The authors claimed good accuracy using their customised evaluation tool called *Complement of the Absolute Normalised Difference (CAND)*, not standard statistical tools (such as correlations); however, we applied this metric to our explorative study result in Appendix B and found that the minimum value of the CAND is varying depending on rates of interest and the output tends to be generally very high (in our case, despite low correlation of $r=0.519$, the CAND-base accuracy was 90.79%, Figure B.2).

pulse measurements (e.g. ECG or PPG), this method is unable to provide a rich set of metrics including Heart Rate Variability (HRV) related metrics (Shaffer & Ginsberg, 2017).

Respiratory thermal signature

Likewise, it has been shown that respiratory rate can be extracted through thermal imaging (Murthy *et al.*, 2004; Murthy & Pavlidis, 2006; Fei & Pavlidis, 2010; Abbas *et al.*, 2011; Lewis *et al.*, 2011; Pereira *et al.*, 2015). This can be done by monitoring thermal patterns of air exchanges inside the nostrils, which are caused by inhalation and exhalation breathing cycles. Although areas around lips and nostrils can be used, most of existing works have used the nostril as a key ROI in which they have averaged temperatures to represent the breathing signal. This approach has been tested within several contexts in indoor, stationary settings where the position of a thermal camera can be fixed, including neonatal care (Abbas *et al.*, 2011) and sleep (Fei & Pavlidis, 2010). Much higher performances have been achieved in the monitoring of breathing cyclic information using thermal imaging than the cardiac pulse monitoring and it has been shown that extracted time-varying signals reflect breathing signature more reliably than other types of physiological thermal signatures (note that the approach capturing cardiac pulse rate does not produce raw blood volume pulse information, and the same applies to the thermal directionality, such as temperature rise/drop, reviewed above). However, this method has also been challenged by difficulties in tracking the nostril.

In the latest work where state-of-the-art performances were achieved (Pereira *et al.*, 2015), the respiration tracking accuracy was greatly improved by using the latest automatic ROI-tracking method in computer vision (Mei & Ling, 2011) within indoor controlled experiments (strong correlation with ground truth: $r=0.974$). This indicates state-of-the-art ROI tracking methods could handle motion artefacts. However, they still do not provide robustness to environmental thermal dynamics. Even though our facial temperature distribution is internally controlled (e.g. blood vessel regulation), ambient temperature affects thermal images. For this reason, the actual temperature within which signals can be found may vary.

Such challenges are similar to the tone-mapping-related quantisation issue found in converting real-world luminance to virtually expressed colour (Ledda *et al.*, 2005). Although lots of challenges have been identified in high dynamic range imaging, existing works using thermal imaging have entirely ignored the issue, adopting a fixed range of temperature (e.g. 28°C to 38°C in Fei & Pavlidis, 2010). This is the component that can deteriorate the tracking performance

under high thermal dynamic scenes. Other than the ROI tracking issue, the quality of the respiratory signature is also influenced by such motion artefacts and environmental thermal dynamics, which has not been broadly investigated by previous authors. Hence, a more robust respiration tracking approach is required to improve the accuracy further in mobile settings.

Muscular thermal signature

Lastly, thermal imaging can also be employed to observe cues related to muscular activations. The activations can closely be associated with behavioural changes. For example, a facial expression is made up of a number of activations of micro-muscle units on the whole face (Ekman, 1993). To perform full-body or partial-body movements, skeletal muscles in our body should be activated (Hall & Guyton, 2015). As for facial expression recognition tasks, earlier, thermography had been used as an auxiliary means to improve performances of RGB vision cameras-based approaches (Heo *et al.*, 2004; Bebis *et al.*, 2006). Later, thermal imaging has become a primary channel for recognising facial expressions (Hernández *et al.*, 2007). This has led to initial investigations of the use of thermal imaging to study affect-related facial muscle activations (Liu & Wang, 2011; Wang *et al.*, 2014b). More details are discussed in Section 2.3.2.

2.3. Thermography as a Measure of Human Affects

With the literature reviewed in Section 2.2, this section aims to bring together the relevant literature that explores the ability of thermal imaging in assessing a person's affective states in order to identify which types of thermal signatures and their metrics have been investigated in relation to affective states. Before starting to review, we would like to draw attention to a diagram presented in Figure 2.2 which summarises existing works in this field (thickness of a line is to show what a majority of existing works have focused on). As discussed above, the selection of ROIs is linked to physiological signature of interest. Amongst explored physiological thermal signatures, vasoconstriction/dilation related cardiovascular activity and perspiratory activity as well as facial muscular activity have been explored in relation to affective states. These works have been conducted in systematic constrained settings in terms of environmental temperature changes and movements. We discuss more details in Sections 2.3.1 and Sections 2.3.2, which are followed by a discussion with summary tables in Section 2.3.3.

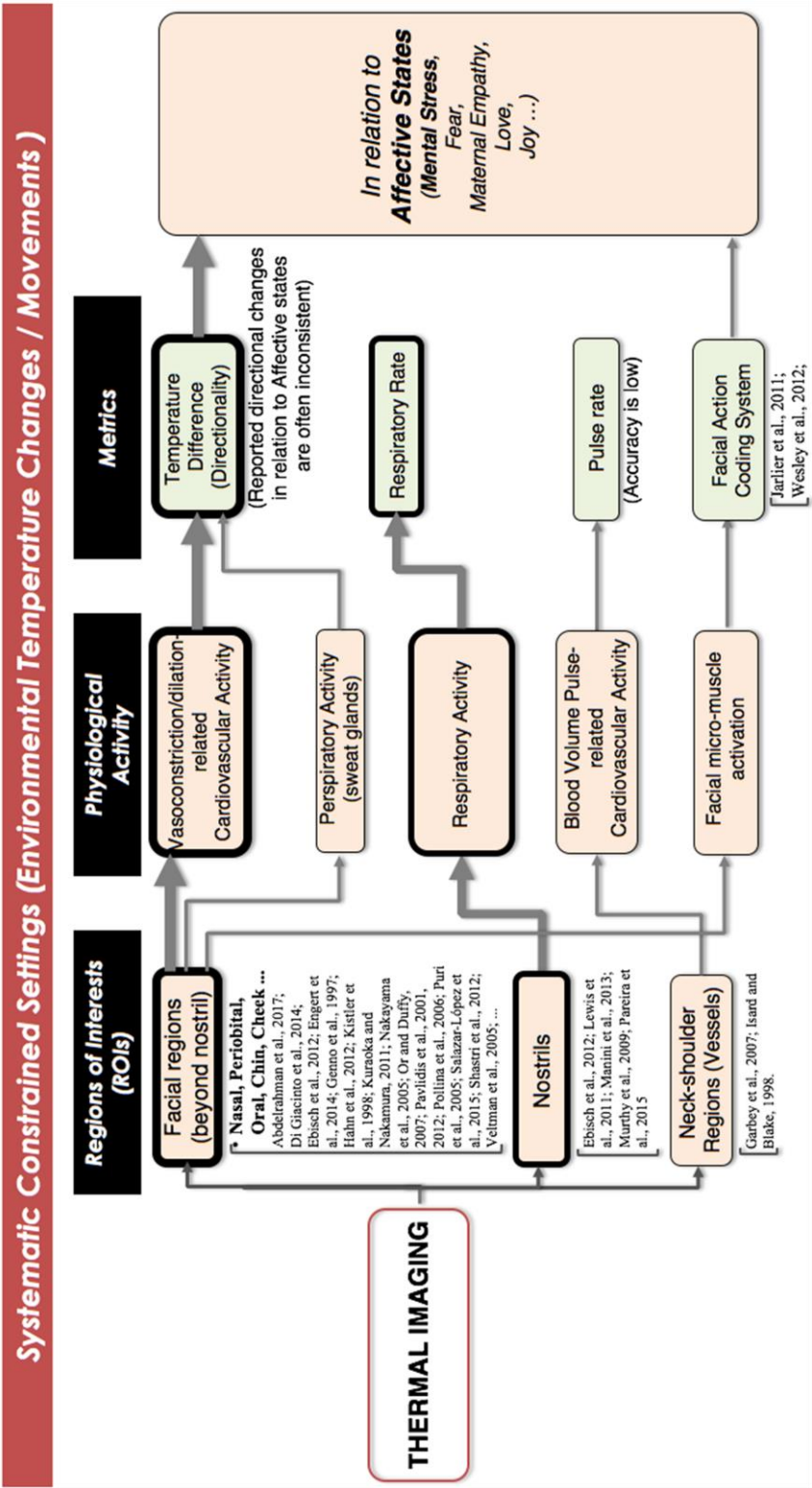


Figure 2.2. A summary of the literature exploring the use of thermography as a measure of physiological cues and affects: choosing right ROIs corresponding to physiological signatures of interest and their metrics are key in assessing a person's affective states (Thickness of a line is to highlight what a majority of existing works have focused on). All existing works have been conducted in systematic constrained settings in terms of environmental temperature changes and movements.

2.3.1. Thermal Directionality in relation to a Person's Affective States

Amongst physiological thermal signatures, the vasoconstriction/dilation-related cardiovascular and sweat gland activation-related perspiratory responses induce increases or decreases in temperatures of ROIs, which could be quantified by using a simple *metric* such as binary directional change (e.g. temperature difference between data at two temporal points of interest). Hence, a majority of studies within this field have focused on thermal directional changes on ROIs in association with a person's affective states. In addition, the body of work have mostly chosen local areas on thermal images of the face (an example of facial thermal image is shown in Figure 2.3 Left). Initial works in this field concentrated on psychological stress or mental workload-induced temperature changes. More recently, interests have been expanded to different types of affective states ranging from anxiety and fear to sexual arousal and maternal empathy in contexts of social interaction. Before diving into existing works, Figure 2.3 shows examples of thermal directional responses of some selected ROIs for peoples affective states. Below are the detailed reviews of the literature.

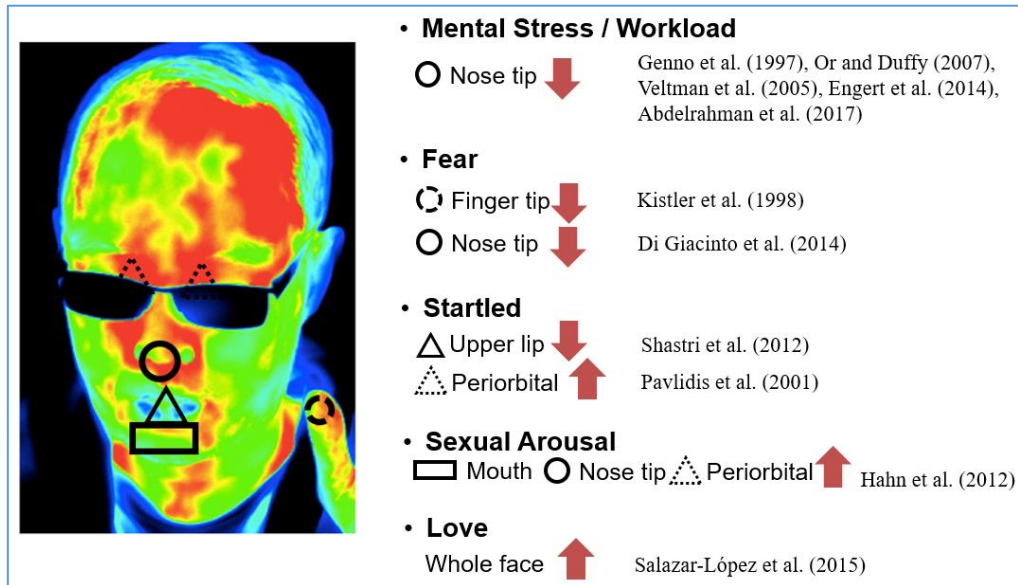


Figure 2.3. Left) an example of thermal image of his face and thumb finger, and widely used ROIs, Right) a quick summary of thermal directional responses of selected ROIs to a person's affective states.

Thermal directional changes in relation to Mental Stress

Genno *et al.* (1997) analysed nose and forehead temperatures to observe their directional responses to mental stressors. Using a contact-based multi-channel thermistor, they measured both temperatures of 117 participants during a resting period and a self-designed stress induction task. During the task, participants were required to keep searching for a moving virtual object on a computer screen which suddenly pops up with a flash lightening effect and a loud siren sound feedback as the onset of stressors. To minimise effects of ambient temperature they controlled the room temperature. In the analysis, they directly compared temperature of the nasal area with that of the forehead region. They reported a significant decrease in nasal temperature under the stress condition. This can be explained by vasoconstriction (i.e. narrowing of blood vessels) of nasal peripheral tissues under the mentally stressed condition (Elam & Wallin, 1987). On the other hand, forehead temperatures of participants were stable. Despite the use of contact-based sensor (not thermal imaging), the result showed greater potentials of the thermal directional change as a stress indicator. This leads to the following studies using different experimental protocols and more importantly, using contact-free high-resolution thermography. ROIs have also been extended to peripheral facial areas including peri-nasal regions, supraorbital and frontalis areas (Puri *et al.*, 2005; Veltman & Vos, 2005; Or & Duffy, 2007; Pavlidis *et al.*, 2012; Engert *et al.*, 2014).

Using a high-resolution thermal imaging system, Or and Duffy (2007) observed significant thermal changes of the nasal area in a negative direction during a driving simulation task (3min) and a real-field driving task (5min) in comparison with pre-driving sessions as the baseline. Both tasks were designed to induce a considerable amount of mental stressors. Or and Duffy confirmed there was no significant change in temperature of the forehead. In addition, authors noted that the Sympathetic Nervous System (SNS)-triggered vasoconstriction could explain the physiological response of the nose temperature to mental workload, which was captured by thermal imaging. Very recently, the same result (significant drop in temperature of the nose tip in comparison with the forehead temperature) was reported from studies using the Stroop colour-word test (Abdelrahman *et al.*, 2017). During the Stroop test (Stroop, 1935), participants are required to name the colour of words while the meaning of each word was presented either incongruent or congruent. This test has been widely used as a means of producing cognitive load-induced mental stress in psychophysiological and neuroscientific studies (Taylor *et al.* 1997).

From another study using the Stroop test as stressors, the role of thermal imaging as a measure of mental stress has also been discussed (Puri *et al.*, 2005). The authors focused on the forehead region, in particular, the skin above a thick blood vessel which is connected from the supraorbital area to the frontalis. They interpreted temperature patterns on the area into the amount of blood volume changes and compared this with the amount of energy expenditures using a gas collection mask (from nostrils and lips), which was used as the ground truth of stress measures (Seematter *et al.*, 2002). Despite details of their computational methods not described in their article they showed a strong correlation between the different measures, indicating that thermal imaging-based cardiovascular measure could support monitoring of psychological stress levels.

As part of an effort to extend ROIs which can show physiological thermal responses to stressors, Engert *et al.* (2014) conducted a study using the trier social stress test (Kirschbaum *et al.*, 1993) which involves a mock-up job interview and a difficult mental arithmetic solving task. The authors observed temperature patterns from periorbital, perioral, corrugator and chin, as well as the nose and forehead regions. In particular, the authors extracted a time series of mean temperatures from each region and also extracted the slope from each timeseries data. This was used as an alternative to selecting two temperatures from the beginning or baseline, and from the end of each session, which is the dominantly used way. For this they visually inspected all thermal imaging data and removed data collected when a participant's head moved. As results, not only was there a decrease in temperature of the nasal area, but also declines in temperature of chin and corrugator areas under the stressful condition. However, from a more highly controlled experiment using a chin rest (Veltman & Vos 2005), such areas other than the nasal area did not respond to mental stressors. During the experiment using a mental memory task, Veltman & Vos (2005) focused on the thermal directional change of 17 facial ROIs including the nose tip, chin and corrugator (outside an eye and around eyebrow). Such incongruent results may indicate limitations of the use of simple directionality-related metrics and the analytic process which depends on manual or visual inspections which may produce experimenter-dependent errors.

In a recent study, Pavlidis *et al.* (2012) focused on perspiratory thermal responses of peripheral areas of the nose and a finger in quantifying mental stress levels, rather than focusing on vasoconstriction/dilation-induced thermal directional changes. Using a very high spatial resolution camera, the authors attempted to directly observe thermal responses of sweat glands to stressors. They used a surgical laparoscopic drill training session as the main stress induction task.

The authors documented that novice surgeons showed higher stress levels than experts, a lower temperature of the perinasal area (above lips and surrounding areas) could be another stress sign.

Thermal directional changes in relation to other affective states spanning from anxiety, fear, to love

Such promising results from the body of work has led to studies where facial thermal directional responses to other affective states have been investigated. For example, the thermal response to anxiety has been explored in the context of deception (Pavlidis *et al.*, 2002; Pollina *et al.*, 2006). Pavlidis *et al.* (2002) conducted an experiment which required participants to commit a mock crime and declare their innocence of the crime. When participants were given a specific forensic question, they tended to produce temperature changes in their periorbital regions. The finding was also reported in studies using different crime-related protocols (Pollina *et al.*, 2006).

The thermal directional response of fingertips to fear-conditions has also been explored (Kistler *et al.*, 1998; Di Giacinto *et al.*, 2014). For example, Kistler *et al.* (1998) used a horror film to induce fear-conditions. Compared with temperature measured prior to the task (watching the film), participants had temperature drops of up to 2°C on their fingertips. In another study conducted by Di Giacinto *et al.* (2014), a sudden acoustic stimulus was used to induce fear conditions in people with Post Traumatic Stress Disorder (PTSD). A temperature drop was generally observed over the whole face of participants. The nose tip showed the highest thermal change (drop). Some studies exploring the facial temperature in nonhuman primates have also supported the evidence for the nasal temperature as a measure of fear (Nakayama *et al.*, 2005; Kuraoka & Nakamura, 2011).

Our facial areas could also have such thermal directional changes along with startle states (Pavlidis *et al.*, 2001; Shastri *et al.*, 2012). Shastri *et al.* (2012) documented that temperature of the upper lip and its surrounding regions of their participants decreased during the presentations of unexpected startling sound effects (e.g. glass breaking sound). Similarly, Pavlidis *et al.* (2001) reported the increase in temperature of the periorbital region of participants was observed after being exposed to a loud startling sound. However, other authors - who conducted a study based on a similar protocol to Pavlidis *et al.* (2001) - reported an incongruent finding (Gane *et al.*, 2011). There were no significant changes in temperature of the periorbital area. The inconsistent results

may indicate that the thermal directional signature can be influenced not only by a variety of internal, however also external factors, for example, environmental temperature.

It is clear that a wide variety of advantages of the non-contact and non-invasive measurement are attractive in its use to observe a person's affective state in social contexts (Ebisch *et al.*, 2012; Hahn *et al.*, 2012; Salazar-López *et al.*, 2015). Hahn *et al.* (2012) studied sexual arousal during social interaction with interpersonal physical contact (i.e. physically touching a body). They identified the closer relationship a participant has with his opposite gender partner, the higher increase of temperature the person showed on the nasal, periorbital areas and the lips. In terms of the socio-emotional development in children, furthermore, Ebisch *et al.* (2012) explored thermal signatures (again, thermal directional change) in relation to empathy between mothers and their children in stressful situations. The authors found a significant parallelism between their facial temperature changes (mother-child). In a recent study, thermal variations associated with *romantic* love in couples were found (Salazar-López *et al.*, 2015). Participants in the study were asked to take a look at pictures of their beloved partner to postulate love in memory and those of their friend during a baseline period. The authors confirmed there was a general increment in their facial temperature in feeling love.

Automated affect recognition

Although those earlier studies highlighted the capability of thermography in affective computing, they focused mostly on observing affect-related thermal directional changes of a certain ROI through manual annotation, not bringing such signatures into automated affect recognition problems. Interestingly, Nhan & Chau (2010) attempted to automatically recognise one's affective states using thermal imaging (the heavyweight, high-resolution system was used). They focused on how to infer a person's arousal (unexcited-excited) and valance (unpleasant-pleasant) states (i.e. high or low). They set five ROIs: left, right supraorbital, left, right periorbital, and nasal, which all are influenced by constriction and dilation of vessels controlled by SNS activity (Wallin, 1981). Participants were given a set of pictures selected from the International Affective Picture System (IAPS) (Lang *et al.*, 2008) in their experiment. With thermography, they also used *additional contact-based physiological sensors* to measure heart rate and respiration rate. To avoid the complexity of automatically tracking each ROI, a dot stick was attached to the top of their participants' forehead. For automated affect recognition, 78 values from each ROI (using basic statistical and customised functions), 10 correlation coefficients

between each ROI pair (i.e. \mathcal{C}_2 – combination) and average heart and respiratory rates (from the contact-based sensors) were used as features. They employed the Linear Discriminant Analysis (LDA)-based machine learning technique as a binary classifier. Their approach achieved success rates of approximately 80% in discriminating high arousal state from a baseline state, and high valence state from a baseline state, respectively. On the other hand, their classifier could not produce above-chance level accuracies in classifying high arousal versus low arousal, and high valence versus low valence.

2.3.2. Facial Muscular Thermal Signature in relation to a Person's Affective States

Facial muscular thermal signature has also been explored in their relation to a person's affective states (Hernández *et al.*, 2007; Wang *et al.*, 2010; Liu & Wang, 2011; Jarlier *et al.*, 2011; Wesley *et al.*, 2012). Facial micro-muscle activations contribute to producing numerous facial expressions. Studies have focused on how to use thermography for automated facial expression recognition (Hernández *et al.*, 2007; Wang *et al.*, 2010; Liu & Wang, 2011; Khan *et al.*, 2016). Similar to RGB-vision based approaches, the body of work attempted to select a couple of (e.g. mostly three to five) facial ROIs which include main sensory organs such as the eyes and lips. By contrast, with studies focusing on the observation of affect-related directional physiological changes of facial temperatures, a variety of machine learning techniques have been used for automatic recognition of facial expressions which appeared on thermal image sequences. For example, Liu and Wang (2011) and Wang *et al.* (2014) extracted features from thermal images, such as a histogram and mean which were applied using Hidden Markov Models (HMMs) to classify expressions relevant to happiness, disgust and fear.

More recently, Khan *et al.* (2016) propose a segmentation of a person's facial thermal image into a number of local areas (e.g. 75 local squared areas) to extract a maximum temperature timeseries from each area. Using a feature selection technique and the LDA, the authors achieved 73.7% and 68.4% accuracies in automatically recognising happiness and sadness, respectively, from their data collection study using IAPS (similarly used in Nhan & Chau, 2010; Salazar-López *et al.*, 2015 – reviewed in Section 2.3.1).

There have been studies which have focused on understanding facial micro-muscle activations which occur in the process of facial expressions. It has been shown that micro-muscle

activations can lead to morphological changes on thermal image sequences (Jarlier *et al.*, 2011; Wesley *et al.*, 2012). Traditionally, the Facial Action Coding System (FACS) and facial ElectroMyoGraphy (EMG) have been used to observe or decode such muscular activations. In particular, the FACS is basically based on the manual analysis of anatomical composition of human facial muscle Action Units (AUs) (Ekman & Friesen, 1977). While this is very accurate, this requires a trained human coder. Hence, researchers have tried to automatize FACS coders' time-consuming decoding works (Jarlier *et al.*, 2011; Wesley *et al.*, 2012).

To the best of the author's knowledge, Jarlier *et al.* (2011) first explored facial thermal patterns associated with the activation of facial AUs during producing emotional expressions. The authors recruited four trained FACS coders. They observed not only the temperature distribution on coders' faces where different action units were intentionally activated to make emotional expressions, but also the speed and intensity of corresponding muscle contractions. The participants' heads were immobilised to minimise motion artefacts. The Principal Component Analysis (PCA) was used to reduce the dimension of features and select informative features of the activated facial action units. For the classification, the k-Nearest Neighbour model (kNN) was used. They reported that amongst eight cases (requesting 8 different AUs activations), the correct rates of six cases were above 60%, showing potentials of thermography in the FACS decoding process. Wesley *et al.* (2012) followed this work by comparing the thermal imaging based facial expression classifier with a RGB vision based classifier. The authors documented that thermography-based approaches could be more useful given the fact that thermography is much less susceptible to the ambient lighting condition issue, which is one of the main limitations in the use of RGB vision-based approaches.

2.3.3. Summary Tables

Table 2.2 summarises existing approaches to thermography-based affective computing, which we have discussed in Section 2.3. The main focus of the literature spans from thermal directional changes of facial local areas in relation to psychological stress, fear, anxiety and to facial muscle activations in association with facial expressions. We elaborate the main information of the study protocols used in the literature (e.g. task types and baseline) in Table 2.2, providing an insight for building experimental protocols. As described in the table, only facial areas have been dominantly explored. In order to quantify the discussed thermal directional responses to affective states,

authors have generally used a simple metric, *the difference between thermal data at two temporal points*. In particular, averaging has been used to represent temperatures of a ROI on thermal images.

Furthermore, existing studies (reviewed in Sections 2.2 and 2.3) have all been conducted in systematic constrained settings in terms of stable environmental (ambient) temperature and low-level motion artefacts. Although a few studies have explored automatic ROI tracking (also called motion tracking) methods, they often reported negative effects of tracking imperfections on affect quantification tasks (Garbey et al., 2007; Sun et al., 2005; Ebisch et al., 2012; Manini et al., 2013; Shastri et al., 2012; Tsiamyrtzis et al., 2007; Zhou et al., 2009). In addition, the tracking performance in varying thermal range scenes has remained unaddressed.

Table 2.2. A summary of thermography-based affective computing studies: experimental protocols (in particular, constraints required), explored ROIs and metrics.

Affective States	Author	Experimental protocols				Analysis	
		Participants	Task	Baseline	Requirements	ROIs	Metrics or Features ¹⁰
Mental Stress	Genno et al. (1997)	117 healthy adults	Self-made tracking task (11mins)	Rest (5 mins)	Stable room temperature	Nasal and forehead	Difference between mean temperature (from two points) & Temperature difference between data from two ROIs
	Veltman and Vos (2005)	8 adults	Continuous Memory Task (3mins)	Rest (3mins)	Chin rest used	Nasal, cheek, chin, lips, eyes, forehead	Difference between mean temperature (from two points)
	Or and Duffy (2007)	33 healthy licensed drivers	Car-driving simulator tests (3mins)	Rest (3mins)	Stable room temperature	Nasal and forehead	Difference between mean temperature (from two points)
	Pavlidis et al. (2012)	17 healthy adults	Laparoscopic drill training (4mins)	Natural landscape	Not reported (but, see Figure 2.1)	Perinasal and Finger	Difference between mean temperature at two temporal points (on sweat glands)
	Engert et al. (2014)	15 male adults	1) Cold Pressor Test, 2) Trier Social Stress Test	Rest (5mins)	Stable room temperature & Visual inspection	Chin, Corrugator, Forehead, Finger, Nose, Peri-orbital, Perioral	Mean slope

¹⁰ Although metrics and features are used as synonym in some literature, in this thesis, *metric* is used to describe a measure to quantify physiological symptoms and *feature* is used to describe specifically the feature input to, or analysed in, machine learning models (in most cases of the thesis, classifiers).

Cognitive Load	Abdelrahman et al. (2017)	24 Egyptian and Canadian participants	Stroop Test	Rest (10mins)	Stable room temperature & Keeping participants facing a thermal camera	Nose tip & forehead	Difference between mean temperature (from two points)
Fear	Kistler et al. (1998)	20 healthy Caucasian adults	Watching horror movie scene (3mins 45s)	Rest (2mins)	Stable room temperature	Finger tip	Difference between mean temperature (from two points)
	Di Giacinto et al. (2014)	10 post traumatic stress disorder patient, 10 healthy adults	Image presentation with 80-db white noise burst (6mins)	Image presentation with no sound (6mins)	Stable room temperature & without ventilation	Nasal and whole face	Difference between mean temperature (from two points)
Anxiety / Deception	Pollina et al. (2006)	13 healthy adults	Committing a mock crime	not specified	Stable room temperature	Right and left hemifaces	Difference between mean temperature (from two points)
Startle reaction	Pavlidis et al. (2001)	6 healthy participants	A startle stimulus using a sudden loud noise (60dB)	Sitting in the dark room (10 mins)	Stable room temperature	Periorbital, Cheeks, Neck, Nasal, chin region	Average pixel values (relative thermal change)
	Shastri et al. (2012)	18 healthy participants	Presentation of unexpected startle sounds (e.g. glass breaking sound).	not specified	Not reported (but, a similar environment to the one in Figure 2.1)	Upper lip and surroundings	Difference between mean temperature (from two points) & wavelet -based metrics
	Gane et al. (2011)	11 adults	102 dB Auditory startle stimulus (instantaneous (30s) x 100 repetitions)	Rest (1min)	Manually adjusting thermal imaging parameters to room temperature (an ambient temperature sensor was used)	Periorbital areas	Range, mean, variance, skewness, kurtosis, entropy, normality, stationarity of 2 second mean temperature timeseries
Sexual arousal	Hahn et al., (2012)	23 healthy adults	Facial contact with an opposite-sex experimenter (15mins)	Prior to facial contact	Keeping participants from moving	Nasal and periorbital areas and lips	Difference between mean temperature (from two points)
Maternal empathy	Ebisch et al. (2012)	12 mothers and their children	Eliciting stressful situations based on the “mishap paradigm”	Prior to task (10-20mins)	Stable room temperature	Nasal tip, maxillary regions	Difference between mean temperature (from two points)
Love	Salazar-López et al. (2015)	12 couples	Presentation of pictures of the beloved partner (5mins)	Presentation of pictures of friends (5mins)	Keeping participants from moving	Whole face, hand	Difference between mean temperature (from two points)

Positive-negative emotions	Nhan & Chau (2010)	12 healthy adults	Presentation of pictures selected from the international affective picture system (IAPS)	Rest	Stable room temperature & a dot tracker used	left, right supra-orbital, periorbital, nasal regions	390 features (5ROIs x 78) and 10 features (correlation coefficients from pairs of ROIs) (*for classification)
	Salazar-López et al. (2015)	120 university students	Presentation of pictures selected from the international affective picture system (IAPS) (several seconds)	Rest (10mins)	Keeping participants from moving	nose, forehead, orofacial area and cheeks and the whole face region	Difference between mean temperature (from two points)
Facial expressions and Facial Action Coding Systems (FACS)	Khan et al. (2004), (2016)	16 undergraduates (2004), 19 participants (2016)	Presentation of pictures selected from the international affective picture system (IAPS) (several seconds)	Rest (20mins)	Stable room temperature	Self-defined facial thermal feature points	Covariance metrics and maximum temperature timeseries from 75 self-selected ROIs on a face (*for classification)
	Wang et al. (2010)	215 healthy students	Presentation of film clips	Not reported	Unconstrained room environments	Forehead, nose, lips, cheek	Feature selection using PCA, PCA+LDA, AAM, AAM+LDA, etc. (*for classification)
	Jarlier et al. (2011)	4 trained and certified FACS coders	Activating different facial action units	Rest (15mins)	Keeping participants' heads immobilised with a head fixation system	Facial action units	Topography of thermal changes based on speeds and intensities of muscle contraction (*for classification)

Table 2.3 summarises the specifications (resolutions, sensitivity) of thermographic systems which have been used in the literature discussed in Section 2.2 and Section 2.3, as well as actual requirements for their thermal analysis. First, all of existing works have employed very heavy and expensive thermographic imaging systems (e.g. Figure 1.1b, Figure 2.1), which are hardly accessible to the general public. In addition, installation difficulty (e.g. top view) has confined ROIs mostly to the facial regions, possibly interfering with one's vision, causing distraction, and more importantly, prohibiting this from being carried out in unconstrained, mobile settings.

Interestingly, many works have required lower spatial resolutions and sampling rates (than specifications of their systems) as summarised in Table 2.3 (right column). For example, very

low sampling rates or static data have been required for capturing thermal directional changes (the sampling rate of 340Hz was down-sampled to 0.2Hz in Veltman and Vos, 2005). Respiratory rate can also be extracted at a lower sampling rate of 5Hz (Lewis *et al.*, 2011). In Shastri *et al.* (2012), the spatial resolution of 640x512 was down-sampled to 100x100.

Table 2.3. Specification requirements for thermography-based physiological measurements and affective computing. The majority of existing works indeed require lower spatial and temporal resolutions than actual resolutions thermal imaging devices have.

Author	Metrics	Affective States	Specification of Thermal Device			Requirement for Analysis		
			Spatial Resolution	Sampling Rate	Thermal Sensitivity	Spatial Resolution	Sampling Rate	Thermal Sensitivity
Fei & Pavlidis, 2010	respiratory rate	Not explored	640x512	55hz	0.025°C	640x512	10Hz	0.025°C
Murthy et al. (2004)			640x512	126Hz	0.025°C	640x512	31Hz	0.025°C
Pereira et al. 2015			1024x768	30hz	0.05°C	1024x768	30Hz	0.05°C
Lewis et al. (2011)	respiratory rate & relative tidal volume		320x240	<29hz	0.08°C	320x240	5Hz	0.08°C
			640x510	126hz	0.02°C	640x510	30Hz	0.02°C
Shastri et al. (2009)	skin-conductance-related raw signal			640x512	126 Hz	0.025°C	640x512	126 Hz
Abdelrahman et al. (2017)	thermal directional change	cognitive load	160x120	120Hz	0.08°C	160x120	120Hz	0.08°C
Di Giacinto et al. (2014)		fear	320x240	60Hz	0.02°C	Not reported	10Hz	0.02°C
Ebisch et al. (2012)		empathy	320x240	50Hz	0.02°C	Not reported	1Hz	0.02°C
Engert et al. (2014)		mental stress	320x240	50Hz	0.02 °C	320x240	5Hz	0.02 °C
Gane et al. (2011)		startled state	640x480	120 Hz	0.01°C	640x480	15Hz	0.01°C
Hahn et al. (2012)		sexual arousal	160x120	1/75 Hz	0.08 °C	160x120	1/75 Hz	0.08 °C
Kistler et al. (1998)		fear	280x90	6.25Hz	Not reported	Not reported	0.1Hz	Not reported

Nakayama <i>et al.</i> (2005), Kuraoka <i>et al.</i> (2011)		fear	Not reported	Not reported	Not reported	254x238	0.1Hz	0.1 °C
Or and Duffy (2007)		mental stress	320x240	60Hz	0.08°C	Not reported	Static Image	0.08 °C
Pollina <i>et al.</i> (2006)		deception-related anxiety	256x256	30Hz	0.1 °C	256x150	30Hz	0.1 °C
Salazar-López <i>et al.</i> (2015)		romantic love	320x240	60Hz	0.07 °C	320x240	Static Image	0.1 °C – 0.5 °C
Shastri <i>et al.</i> (2012)		startled state	640x512	126 Hz	0.025°C	100x100	25Hz	0.025°C
Veltman and Vos (2005)		mental stress	320x256	340Hz	0.07°C	320x240	0.2Hz (every 5s)	0.07°C
Pavlidis <i>et al.</i> (2012)	<i>thermal directional change & skin-conductance-related raw signal</i>	mental stress	640x512	126Hz	0.025°C	640x512	25Hz	0.025 °C

As described in Figure 2.2 which summarises thermography-based studies for physiological measurements (Section 2.2) and for affective computing (Section 2.3), it is noteworthy that while the thermal directional changes of facial regions have been dominantly investigated so far, other types of physiological thermal signatures and body parts beyond the face have been underexplored in this body of work. Given this, in the next section, beyond the primarily focused facial areas, we elaborate physiological evidence to find potential ROIs, in particular which can provide information of a person's stress.

2.4. The Scalp as a Potential Region of Interest (ROI) for Stress-related Thermal Signatures: Neuroscientific Evidence

Although identifying facial thermal signatures associated with stress has been one of the most dominant topics in this field of study, signatures from various body regions have been

underexplored. The scope of this review is confined to neurophysiological responses of the brain to psychological stressors and their thermal signatures (Additionally, the physical stress responses of skeletal muscles as another facet of stress are reviewed in *Appendix A* to provide more insights).

The brain is a human pivotal organ which processes cognitive and psychological activities. It controls physiological and behavioural responses to numerous intrinsic and extrinsic factors. In particular, against mental stressors, the organ arbitrates neurological mechanisms which cause a series of physiological, metabolic, and behavioural responses. Below is the review to find neurological and physiological evidence of the heat production or alteration in the brain under mental workload. This leads to a discussion of searching novel ROIs where stress-related thermal signatures can be potentially captured through thermal imaging.

2.4.1. Stress Responses of the Brain in relation to Temperature

Hippocampus, hypothalamus, prefrontal cortex (PFC), and amygdala are major cerebral regions which react to mental stress (Dedovic *et al.*, 2009). The regions are in charge of controlling chemical releases and activation/deactivation of local functional areas of the brain under mental workload. Here we discuss the mental stress-induced neurological phenomena in terms of *1) releases of biochemistry such as Corticotropin-Releasing Hormone (CRH)* (Dunn & Berridge, 1990; Koob *et al.*, 1993) and *Dopamine* (Pruessner *et al.*, 2004; Belujon & Grace, 2015), *2) regional changes in Cerebral Blood Flow (CBF)* (Soufer *et al.*, 1998; Pruessner *et al.*, 2004; McEwen *et al.*, 2015).

First, Corticotropin-Releasing Hormone (CRH), also known as Corticotropin-Releasing Factor (CRF), is a neurotransmitter in the mammalian Central Nervous System (CNS). The releasing mechanism of the CRF is closely related to mental stressors (Dunn & Berridge, 1990; Koob *et al.*, 1993). In the view of the hypothalamus-pituitary-adrenal (HPA) axis which engages in the CRH releasing mechanism and triggers the fight-or-flight responses (Kudielka *et al.*, 2004; Pariante & Lightman, 2008; Belujon & Grace, 2015), the CRH is released by the hypothalamus responding to mental stressors, and the hormone manages the secretion of adrenocorticotrophic hormone (ACTH) and glucocorticoids, which influence the central nervous, metabolic, and cardiovascular, respiratory activities (Dedovic *et al.*, 2009; McEwen *et al.*, 2015). Likewise, it has been shown that the dopamine (DA) system engages in the neurological reactions to stressors, releasing dopamine (Pruessner *et al.*, 2004; Belujon & Grace, 2015). In particular, Pruessner *et*

al. (2004) identified a mesolimbic dopamine release which is triggered by psychosocial stressors. The amount of the released hormone can be predicted by monitoring the reduction of [^{11}C] raclopride binding potential (BP) which is proportional to the cortisol level influenced by the ACTH.

Despite a lack of studies exploring whether such biochemical signals directly trigger temperature variances in a person's brain or body, there have been studies exploring their relations in animals, for example, the CRH elevating core temperature in rats (Heinrichs *et al.*, 2001). In addition, psychiatrists documented, "*in humans the rise in body temperature prior to a stressful event may be considered as an index of involvement (primary or secondary) of the anterior hypothalamus and of the corticotropin-releasing factor.*", after they conducted a study where they found an increase in the bodily temperature of a healthy person who felt stressed before a stressful examination period (Marazziti *et al.*, 1992). From the study of long-term consequences of stressors (several months), they concluded by suggesting that, together with blood pressure and rate, the temperature could be another companion of mental stress.

Second, compared with CRH and dopamine releases, **CBF**, the blood supply to the brain in a given period of time is more strongly related to temperature of the brain and is helpful to localise its activated/deactivated regions thanks to brain imaging equipment such as functional Magnetic Resonance Imaging (fMRI). CBF is a homeostatic parameter strongly associated with a cerebral heat production (Kiyatkin, 2010; Wang *et al.*, 2014a; Li *et al.*, 2017). In particular, CBF plays a key role in maintaining brain and body temperature coupling, and the variances in local, sometimes called regional, CBF can be used as a measure of functional brain metabolism which is associated with the brain circulation and core temperature (Govier *et al.*, 1984). Furthermore, the body of work which studies the effect of task difficulty on the regional CBF has shown that some regional CBF would be higher during a stressful task rather than during rest-periods (Weinberger DR *et al.*, 1986; Gur *et al.*, 1988), potentially elevating the corresponding regional temperature.

2.4.2. Potential Local Scalp Regions providing Cortical Thermal Signatures

As reviewed in Section 2.1, in 1878, Lombard reported a rise in temperature from three segmented scalp areas (anterior, middle, posterior) in association with cognitive and emotional processes. Ever since this earlier discovery, investigation on the scalp temperature and emotion

has been being inactive for centuries. In order to discover potential ROIs which could propagate thermal signatures, we elaborate studies on regional changes in CBF under mental workload or mental stressors, given the strong association between regional CBF and core temperature (Li *et al.*, 2017).

Soufer *et al.* (1998) undertook an experiment to find in which cortex areas CBF increases or decreases under the mentally stressful conditions elicited by an arithmetic serial subtraction task using Positron Emission Tomography (PET) equipment and Statistical Parametric Mapping (SPM) analysis. The authors identified that there was a significant effect of such stressors on the blood activation in the *left frontal gyrus* of healthy participants, which is situated nearby the temporalis muscles on the human head (Soufer *et al.*, 1998). In the case of patients with coronary artery disease, they showed even more fluctuated patterns of CBF on several different regions. In addition, Taylor *et al.* (1997) also unearthed activations and deactivations of regional CBF in response to mental stressors in a study where the Stroop test was used; for example, the *left inferior frontal gyrus* from the frontal lobe, the regions across the *parietal lobe* (more specifically, *left parietal/occipital junction*, *right inferior parietal lobules*) were activated and the *superior temporal gyrus* was deactivated. In this study, PET was used (Taylor *et al.*, 1997).

With the growing importance of fMRI, a study using fMRI showed the activation of CBF in the *right prefrontal cortex* of a person facing psychological stressors (Wang *et al.*, 2005). The arithmetic task-based protocol was used in the study. Meanwhile, neuroscientists and psychologists have started to discuss the limited amount of stressors in such study protocols in brain imaging environments even though mental stress can be induced by increasing cognitive load from arithmetic solving or Stroop-based tasks (Dedovic *et al.*, 2005, 2009; Pruessner *et al.*, 2008). Given this discussion, Dedovic *et al.* (2005) proposed a new protocol built on the arithmetic solving, called the Montreal Imaging Stress Task (MIST) which adds social evaluative threats within itself. According to the protocol, an experimenter is required to inform participants of average performance of others (using manipulated scores to artificially stress their poor performance). Based on this protocol, Pruessner *et al.* (2008) explored stress-induced regional CBF patterns using both PET and fMRI. The authors identified regional CBF changes in the limbic system as reactions to the mental stressors. In particular, the *left inferior frontal gyrus*, *occipital lobe*, *left prefrontal cortex* and *left premotor* were activated, and the *medial prefrontal cortex* and *superior temporal gyrus* were deactivated in response to the stressors.

More recently, functional near infrared spectroscopy (fNIRS), a relatively new measure of CBF (which measures the oxygenated *HbO2* and deoxygenated *HHb* haemoglobin in the blood

supply in the brain) has been used in this field of study (Al-Shargie *et al.*, 2016; Causse *et al.*, 2017; Rosenbaum *et al.*, 2018). Taking advantage of better portability of fNIRS than PET or fMRI, Causse *et al.* (2017) observed CBF patterns in the brain of participants who took a stressful aviation task using a flight simulator. The authors identified CBF activation across the *prefrontal cortex*. Amongst local *prefrontal cortex* regions, they found the *right dorsolateral prefrontal cortex* was activated most. Al-Shargie *et al.* (2016) also reported the strong correlation between CBF in the right prefrontal cortex and subjective self-report of mental stress scores.

The findings are invaluable for this thesis which found reported areas are closely situated to the skull. This could propagate thermal energy to nearby scalp regions. Hence, there is a high possibility that the mentally distressed conditions could induce thermal variations on the scalp. Other than such cortical regions, part of the central cortex responding to stressors, such as the hippocampus and hypothalamus, are less likely to directly propagate thermal energies to the scalp areas because of their position.

Lastly, in order to shine on the explorations of thermal signatures in brain areas, it should be discussed how to monitor and estimate the brain regional temperature noninvasively. Obviously, it is difficult to measure the core or inner skull temperature of human participants *in-vivo*, for example, while doing everyday desk activities which do not allow to set a huge measurement device (e.g. Magnetic Resonance Spectroscopy). However, physiologists identified that there was a strong correlation between temperature of the temporalis muscle and the cortical temperature, brain core temperature, of rats (Busto *et al.*, 1987). Furthermore, fortunately, a very recent study showed a strong correlation between CBF and brain core temperature of rats (Li *et al.*, 2017), indicating that CBF activation and deactivation patterns could be estimated from the scalp temperatures. As it has been emphasised that there were great similarities in vascular networks of the brains between human beings and rats, so the findings even from rats could be valuable in estimating temperatures of local human brain cortical areas.

Finally, it is certain that the psychological stress severely affects human brain mechanisms as well as physical and psychological conditions on an everyday basis. Moreover, it stimulates the brain's genetic changes (Arnsten, 2009). For instance, the PFC, a part of the brain region which has the highest-order cognitive abilities, is influenced both by short-term and long-term stress exposures, which causes a loss of prefrontal cognitive abilities and architectural changes in the region, respectively. Therefore, it is required to improve our understanding of signals from the brain region and its responses to stressors, and to seek for non-invasive technical intervention

methods mediating mental stress in daily life. In this thesis, thermal signatures on the scalp region are thoroughly investigated for this purpose.

2.5. Chapter Summary and Problem Statements

Existing works in areas of thermography-based physiological, affective computing as well as human physiology on heat production have been discussed in this chapter. Given that thermography raises less privacy related issues and can help remotely read stress-related physiological thermal signatures from a person, not affected by illumination, it is evident that thermography has greater potentials in its use as a multimodal physiological, stress sensor for real-world settings. With advanced thermal imaging technologies available, thermography has been employed in a variety of applications (e.g. medical diagnosis). An overview of their use cases and commercial thermographic systems has been described in Section 2.1.

In Section 2.2, the capability of thermography in physiology measurements was discussed. We found that several physiological signatures could possibly be extracted through thermography. They are: cardiovascular, respiratory and perspiratory thermal signatures mostly from facial areas. In particular, vasoconstriction, vasodilation (cardiovascular) and sweat gland activation (perspiratory) lead to a directional change in temperature (temperature rise or drop). As this can be more easily obtained than other signatures and their metrics, the majority of studies about thermography-based affective computing have focused on which facial areas show an increase/decrease in temperature in association with affective states. This has been discussed in Section 2.3. Affective states explored in the body of work span from mental stress, fear, startle, to love and maternal empathy.

Lastly, as part of an effort to extend stress-related thermal signatures, we have reviewed human neurology and physiology in heat production from the scalp which has been underexplored in Section 2.4.

Through the literature review, we have identified key limitations of existing approaches in building thermography-based physiological computing and stress recognition systems that can support more real world-like settings. Below are further details:

i) Lack of computational methods supporting thermal imaging in unconstrained, mobile situations

Up to this point, computational methods for thermal image processing have been limited mainly due to existing studies all conducted in highly constrained, systematic experimental settings using high-cost, heavyweight thermographic systems (heavyweight as shown in Figure 2.1). In particular, room temperatures and participants' movements have been generally highly controlled in the literature (e.g. stable ambient temperature, often chin rest used) so as to obtain reliable results. To support unconstrained and mobile situations, such constraints should be tackled. At the same time, extractions of thermal signatures need to be automatized. With a concern about motion artefacts, a couple of authors have applied automatic ROI tracking methods (originally built for RGB vision images) to thermal images (despite still limited head movements involved in state-of-the-arts, e.g. Pereira *et al.* 2015). However, effects of environmental temperature changes have never been addressed.

Bringing thermal imaging into more unconstrained settings, it is expected that variable thermal range scenes could affect performances of existing state-of-the-art methods in extracting physiological thermal signatures. As discussed in Section 2.2, thermal dynamics (i.e. different ambient temperature and its dynamic variation) could change the morphological and graphical properties in thermogram and lead to tracking imperfections. In this view point, new signal processing techniques are required to overcome such limitations.

ii) Challenges in tracking good quality of physiological thermal signatures

Amongst explored physiological thermal signatures, vasoconstriction/dilation-related cardiovascular and respiratory signatures have been most often explored. They are very interesting as they are strongly related to mental stress responses (Everly Jr & Lating, 2012) and at the same time, they do not necessarily require high-resolution thermal imaging systems (Table 2.3). By contrast, the perspiratory signature which is also an important physiological stress cue (Healey & Picard, 2005) requires use of extremely high spatial thermal cameras for the monitoring of sweat glands (Pavlidis *et al.*, 2012; Krzywicki *et al.* 2014). In the case of cardiac pulse-related signature which is another type of cardiovascular thermal cues, the accuracy has been very low as discussed in Section 2.2 (and Appendix B).

However, two challenges have emerged from the literature about tracking ROIs containing the two interesting thermal signatures. First, connected to the discussion point above, existing

state-of-the-arts can only support constrained environments in terms of motion artefacts and environmental temperatures. In particular, the prior arts for monitoring vasoconstriction/dilation related thermal phenomenon from the nasal area do not support to continuously and reliably track the area and the signature. Although state-of-the-arts for continuous respiration tracking have reached good performances in tracking the nostril (the main ROI), they still require such constrained environments.

Second, there is a need to improve quality of such physiological thermal signatures. This is directly linked to metrics which can be used to quantify the signatures. Although (general) vasoconstriction/dilation has been known to be very informative of mentally stressed states (Elam & Wallin, 1987), only a simple metric (i.e. binary direction) has been applied to thermal changes in relation to the signature. Continuously extracting higher quality signals should contribute to capturing more complex physiological phenomenon informative of mental stress. The same applies to respiratory signature. Rather than producing a simple breathing rate over a certain time period, there is the need for reliably recovering time-varying breathing-induced air exchange patterns.

Beyond documented (explored) thermal signatures which have been mostly observed from facial areas, we have found possibilities of capturing novel types of thermal signatures from underexplored scalp local regions. From the literature reviewed in Section 2.4, we have found there is sufficient evidence to support the relation of cortical thermal responses of the scalp local regions (e.g. near parietal lobes and inferior frontal gyrus) to mental stressors. We expect mobile thermal imaging can be of greater help to observe such potentially important signatures from various underexplored areas.

iii) Limited contributions of stress-related thermal signatures to automated stress recognition

As addressed above, a variety of existing studies have focused on the simple metric designed to observe binary directions of thermal changes (i.e. temperature decrease, increase) in association with a person's affective states. Despite findings in the literature opening up exciting prospects for thermography-based affective computing, they have often shown incongruent results of thermal directional patterns in association with psychological states (e.g. Stress: Engert *et al.*, 2014 - Veltman & Vos, 2005; Startle: Gane *et al.*, 2011 - Pavlidis *et al.*, 2001). Given our complex psychophysiological mechanisms related to mental stress, a single discrete metric itself

(e.g. thermal directional changes of the nose tip, Genno *et al.*, 1997; Engert *et al.*, 2014) is less likely to contribute strongly to stress inference tasks. The same generally applies to other physiological signatures including respiratory signature.

As summarised in Figure 2.2, the ability of respiratory thermal signature in stress assessment has never been explored. Although, beyond the literature focusing on thermography, gross statistical features (e.g. mean) of respiratory rate have been explored for automatic stress recognition, the contribution of respiratory rate to the stress assessment process has been known to be very weak (McDuff *et al.*, 2016). This could indicate that currently used computational methods should be tackled by building new ways to represent such signatures as respiration is known to strongly respond to mental stressors in the literature in physiology, medicine (Grossman, 1983; Everly Jr & Lating, 2012).

Lastly, whilst the majority of existing works about stress assessments require long-term measurements, for example, longer than five minutes and additional baseline measurements (Engert *et al.*, 2014). However, it is hard to ensure to obtain reliable physiological signatures during this long period of time in ubiquitous situations. Hence, the power of short-term measurements needs to be addressed.

In summary, little has been known about how thermal signatures contribute to building automated stress recognition systems, in particular, supporting unconstrained settings (regardless of the modality). Relevant libraries of computational methods and datasets of thermal video recordings have also remained publicly unavailable. To foster this promising work in our research communities, there is a need for making them available to the public. Thus, this thesis aims to address such challenges.

Chapter 3

Research Methodology

Through this thesis, we define **Mobile Thermal Imaging** as a new pathway for thermal imaging that can perform in unconstrained, mobile, real-world like situations, differentiating itself from the traditional use of thermal imaging within constrained, systematic environments, as shown in Figure 3.1.

In this chapter, we discuss our research questions introduced in Section 1.2 with respect to findings and limitations of the literature reviewed in Chapter 2. Along with this, we present our general approach used to tackle the research questions, materials and tools developed to support our approach.

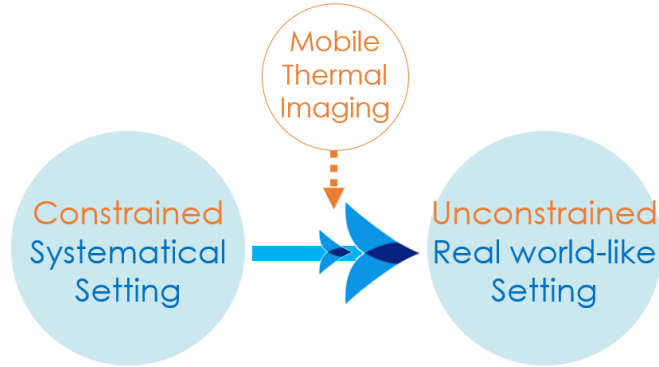


Figure 3.1. Mobile thermal imaging as a new way for bringing thermal imaging from traditional constrained systematic settings into unconstrained, mobile, real-world like situations.

3.1. Three Stage Approach

To address the identified three key limitations, the methodology proposed for this thesis is organised in three corresponding stages: *STAGE I*: Novel Signal Processing Techniques for Mobile Thermal Imaging, *STAGE II*: Novel Physiological Computing Techniques, *STAGE III*:

Novel Automatic Mental Stress Recognition Systems. The following is dedicated to the detailed description of the three stages which form a backbone of this thesis in Table 3.1.

Table 3.1. Proposed three stage approach based upon the research question tree in Figure 1.2.

“Can mobile, low-cost thermography be used as multiple physiological measures for automatic recognition of a person’s mental stress in HCI settings?”

with incrementally growing datasets collected from both controlled settings and real-world-like unconstrained settings.

Evaluation

Stage I

Novel Signal Processing Techniques for Mobile Thermal Imaging

CHAPTER 4	
Novel Pre-processing Techniques for Mobile Thermal Imaging in Variable Thermal Range Scenes	
Rationale The effect of thermal environmental changes, one of the key challenges of thermal imaging in unconstrained settings, has been unaddressed.	
Automatic (Skin) ROI Tracking Experiment Tracking two ROIs on data from Unconstrained & Indoor sedentary task (N=10)	Generalising to other material types Experiment Automatic Material Classification of 26,584 labelled thermal images from 17 outdoor materials

Stage II

Novel Physiological Computing Techniques

CHAPTER 5	
Robust Respiration Tracking	
Rational Need for extreme robustness in breathing tracking in unconstrained settings	
Experiment Tracking the nostril on, and extracting respiration signals from data collected during Controlled Respiration Exercise task (N=5), Unconstrained & Indoor Sedentary Activity task (N=10), Unconstrained & Outdoor Physical Activity task (N=8)	
Ground Truth Signals from a respiration belt	
CHAPTER 6	
Robust Tracking of Stress-induced Nasal Thermal Variability	
Rational Need for robustness in tracking thermal variability from the nose tip area in less constrained settings	
Experiment Tracking stress-induced nasal thermal variances on data from Unconstrained Sedentary Stress Induction task (N=12)	
Ground Truth Task Type & Subjective Self-reports of mental stress	
CHAPTER 7	
Exploration of Novel Cortical Signature	
Rational Cortical thermal signatures in association with stress have been unaddressed	
Experiment Extracting thermal signatures from the scalp and statistically testing extracted data from Unconstrained Sedentary Stress Induction Task (N=10), Highly-constrained Task (N=10)	
Ground Truth Task Type & Subjective Self-reports of mental stress	

Stage III

Novel Automatic Mental Stress Recognition Systems

CHAPTER 8	CHAPTER 9
Representing physiological thermal signatures and Inferring a person's perceived mental stress levels.	
Deep Learning of Respiration Dynamics Rational Unclear contribution of respiration to stress recognition Experiment Stress inference on data from Unconstrained Stress Induction Tasks (N=8) Ground Truth Subjective Ratings	Combining thermal imaging with PPG Rational Need for instant stress recognition Experiment Stress inference on data from Unconstrained Stress Induction Tasks (N=17) with 20second instant measurements Ground Truth Subjective Ratings

STAGE I. Novel Signal Processing Techniques for Mobile Thermal Imaging

The first stage (Chapter 4) aims to answer the research question, RQ1.A) *How can we address the effect of variable thermal range scenes, which is one of the key challenges in automatic tracking of ROIs (Regions of Interest) in real world, mobile situations?* As discussed in Section 2.5, a key rationale for this stage is that the effect of thermal environmental changes, one of the key challenges in keeping track of ROIs on thermal images in unconstrained settings, has been unaddressed in the literature.

To overcome this limitation, this stage focuses on contributing pre-processing techniques to support automatic ROI tracking in unconstrained settings with low-cost, mobile thermal cameras, building on state-of-the-art motion tracking algorithms (originally built for RGB visions) (e.g. Kalal *et al.*'s Median Flow, 2010). In particular, we focus on building a technique which can make tracking performance less affected by environmental thermal dynamics. By doing so, we tackle the standard quantisation process using a fixed thermal range of interest (e.g. 28°C to 38°C in Fei & Pavlidis, 2010) to approximate measured temperatures into grey-colour-scaled values. This standard approach cannot address time-varying environmental temperatures affecting graphical and morphological structures of skins or objects on thermal images, leading to difficulties in tracking ROIs. Hence, we aim to develop a new technique that can adaptively construct a colour mapping of absolute temperature to improve segmentation, classification and tracking.

Furthermore, this stage aims to answer the second research question related to the first one, RQ1.B) *How can we generalise the proposed signal processing approach to other material temperatures than human skin temperatures?* Although the main focus of this thesis is on physiological and psychological cues derived from the skin temperature, we expect that a technique which adaptively adjusts thermal images to environmental temperature changes can be generalised to other types of materials (which have higher thermal conductivities than the human skin). Hence, beyond the problem of the skin ROI tracking, we aim to test this further on thermal images of other materials, particularly, collected from outdoors where environmental temperatures are highly dynamic. By doing so, it is expected to see how important a pre-processing step is for mobile thermal imaging in other HCI applications where thermal imaging is possibly used.

STAGE II. Novel Physiological Computing Techniques

In this second stage (Chapters 5, 6 and 7), we dive into different physiological thermal signatures: respiratory, cardiovascular, and cortical thermal signatures. The first two are the documented physiological signatures in the literature where there is need to improve signature tracking performance, and the third is a novel signature backed by neuroscientific evidence reviewed in Section 2.4.

- 1) Respiratory thermal signature: While earlier works have allowed for small amounts of head motion, most works exploring respiratory thermal signature have used static contexts in which people try not to move (Murthy & Pavlidis, 2006; Fei & Pavlidis, 2010; Abbas *et al.*, 2011; Lewis *et al.*, 2011; Pereira *et al.*, 2015). This leads to the research question, Q2.A) *Can respiratory signatures be recovered through thermography in unconstrained HCI settings? If so, how can we build methods for robust respiration tracking using thermal imaging?*

With the effect of variable thermal range scenes, the first stage is to address the remaining key challenge which has emerged from the literature. The challenge is the effect of motion artefacts to which nostril tracking on thermal images is susceptible. It is also expected that the low spatial resolution of the imaging and mobile situations introduce noise in the respiratory signal quality. In this case, averaging which has been dominantly used to represent respiration signatures (Murthy & Pavlidis, 2006; Fei & Pavlidis, 2010; Abbas *et al.*, 2011; Lewis *et al.*, 2011) is more likely to lose important breathing information. Hence, in Chapter 5, we aim to overcome both issues. Accordingly, we investigate (i) how to enhance the boundary around the nostril region, contributing to making the system more robust to motion artefacts, as well as (ii) how to benefit from the two-dimensional data taken from thermal imaging to contribute towards higher quality breathing patterns.

- 2) Nasal Vasoconstriction/dilation-related Cardiovascular thermal signature: Measuring binary directional changes in temperature of the nasal area has been the main focus of the literature (Genno *et al.* 1997; Veltman & Vos, 2005; Or & Duffy, 2007; Engert *et al.*, 2014; Abdelrahman *et al.*, 2017). As this measurement is of relative ease, state-of-the-art methods to track this physiological phenomenon through thermal imaging have been very limited. As vasoconstriction/dilation itself has shown to be

informative of a person's stressed states in the physiological literature (Elam & Wallin, 1987), we are interested in complex variations in vasoconstriction-related nasal temperature. This leads to the research question, RQ2.B) *Can we continuously monitor stress-induced vasoconstriction/dilation patterns from the nose tip in unconstrained settings? And can we build a rich set of metrics to quantify variations in the patterns?*

Accordingly, Chapter 6 aims to build a novel technique that continuously and accurately estimates vasoconstriction/dilation-related thermal signatures in less constrained settings. This can be based on the outcome from the previous chapters. Due to the low spatial resolution of the thermal camera, the shape of the nose tip region is likely to be blurred, inducing the difficulty in tracking. Hence, we investigate the use of a larger ROI including the nose tip and the effect of breathing induced temperature changes on the measures.

Continuously tracked signals make it possible to extract richer information related to mental stress than the single metric (i.e. binary direction). To support this, we design multiple metrics and evaluate them to see if they lead to better understandings of stress related factors such cardiovascular events from the nasal area. Using datasets we collect from stress-induction tasks (details are given in Section 3.3 and each chapter), we explore the relation of the signature to mental stress further.

- 3) Cortical thermal signature: Beyond known thermal signatures which are mostly from a person's facial area, a potentially interesting signature has emerged from a review of neuroscientific studies in Section 2.4. This is the cortical thermal signature from the scalp. Given neurological evidence of the relation between mental stress and brain activities that potentially influence the scalp temperature (Busto *et al.*, 1987; Taylor *et al.*, 1997; Soufer *et al.*, 1998; Li *et al.*, 2017), Chapter 7 aims to answer the research question, RQ2.C) *Can thermal responses of brain local regions to mental stress be observed by using low-cost thermal cameras? If so, how can we build methods and metrics to capture cortical thermal signatures?*

For this, we can take advantage of the lightweight, small-sized characteristics of low-cost thermal cameras in observing cortical thermal signatures from the scalp. The low-cost, mobile thermal cameras can be set up over a person's head easily, by contrast with the high-cost, immobile thermal cameras. This recording setup does not

require obstruction of vision in real life situations. With this setup, we investigate thermal signatures from brain-scalp regions in relation to Cerebral Blood Flow (CBF) reaction under stressful conditions used in stress induction studies. As in the second part of this stage, we build a richer set of metrics beyond the single and simple direction-related metric to capture stress-induced cortical thermal variations.

STAGE III. Novel Automatic Mental Stress Recognition Systems

In this third stage (Chapters 8 and 9), we explore how to improve the ability to detect mental stress with each physiological thermal signature. Particularly, we focus on respiratory and cardiovascular signatures as the exploration of cortical thermal signature is on a very initial stage. The computing literature has shown that roles of such signatures have been weak in the stress inference process and incongruent results from the use of simple metrics to quantify mental stress have often been reported. That is why we investigate how to address the issues by focusing on each signature separately rather than combining all signatures together to obtain better performance. By doing so, we throw the first research question, RQ3.A) *How can respiratory thermal signature itself contribute to building automatic stress recognition systems?* To answer this, in Chapter 8 we tackle the use of the dominantly-used gross statistical, hand-engineered features from physiological signatures given their limitations. Inspired by a physiologist's discussion of mental stressors possibly inducing breathing irregularity (Grossman, 1983), we aim to focus on how to capture a richness of stress-induced breathing dynamical information. This leads us to build new physiological signal representation techniques that help to condense important stress-related respiration variability. This new compact image-based representation can enable the use of automatic feature learning algorithms (representation learning) such as a Convolutional Neural Network, which could learn informative features itself from the condensed information. Rather than building new machine learning algorithms or making edits on hyperparameters or architectures, we aim to explore how a representation technique can enhance the stress inference performance of existing machine learning classifiers.

The second research question in this stage RQ3. B) is *How can the nasal vasoconstriction/dilation related thermal signatures contribute to very fast stress recognition together with other types of cardiovascular signals from another mobile sensing channel (HRV from PPG)?* Current state-of-the-art stress recognition systems require long-term reliable measurements (e.g. two minutes in McDuff *et al.*, 2016) to make inference about a person's mental stress. However, it is often hard to ensure reliable physiological signatures are obtained

during this long-term period in unconstrained, mobile, and real-world-like situations. Hence, Chapter 9 focuses on improving the ability in automatically detecting mental stress levels using a short-term measurement. In this case, we take advantage of the use of a smartphone and a low-cost thermal camera. This recording setup allows for simultaneously recording two different cardiovascular signatures together: vasoconstriction/ dilation related signatures from the thermal camera and blood volume pulse information from an RGB camera equipped in the phone. Again, the focus is on proposing new representations capturing complex information of mental stress from such signatures. Overall, we aim to achieve state-of-the-art performance in detecting mental stress levels through mobile thermal imaging.

3.2. Basic Tools for Thermal Imaging in Unconstrained Settings

For mobile thermal imaging in unconstrained settings, we use low-cost thermal cameras as thermal imaging sensing channels due to their small-sized, lightweight and portable characteristics making them more feasible to be deployed in real-world settings.

Low-cost Thermal Camera as a Mobile Thermal Imaging Channel

Amongst low-cost, mobile thermal cameras commercially available, FLIR One (for Android - dimensions: 72mm x 26mm x 18mm, FLIR Systems Inc., Santa Barbara, CA, www.flir.com), which was cheapest in 2016 (see Table 2.1), was chosen as a main channel for supporting mobile thermal imaging throughout this thesis. This device converts 8-14 μ m radiation spectral density to temperature. As shown in Figure 3.2, this device requires a smartphone to be connected to control the parameters for the temperature measurement process and to visualise, and store, thermal images. The camera can be either directly connected to the phone or connected to it with an extension cable. When attached to a smartphone, it can be used to continuously sense and record thermal image sequences with a spatial resolution of 160x120 and a temporal resolution of less than 9 frames per second (fps). The device can be placed on a desk or handheld by a user to measure, and record temperatures of a person's face and body, providing more flexibility than high cost, heavyweight thermographic systems which have been dominantly used in the literature (e.g. Figure 2.1).

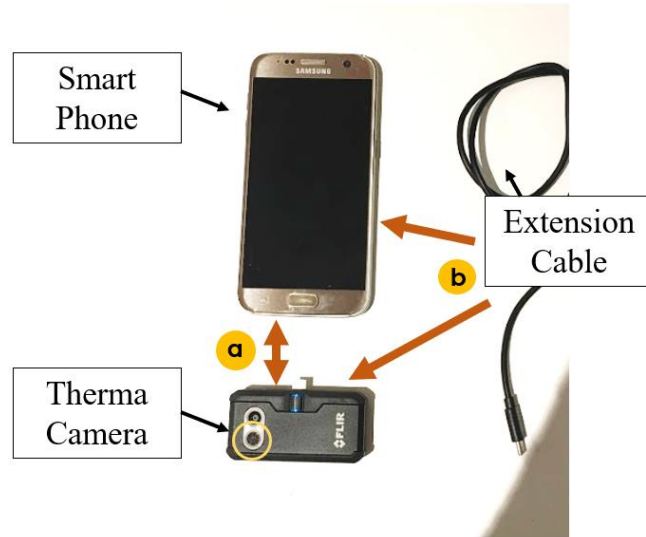


Figure 3.2. A low-cost thermal camera (FLIR One) to be used as a mobile thermal imaging channel in the thesis. This device requires a smartphone to be connected to control parameters for itself and to visualise, and store, thermal images. They can be either directly connected or connected with an extension cable.

Built Bespoke Software Library for Thermal Data Acquisition

An acquisition of measured temperatures (i.e. raw/original data) is the first step to be done for the thermal analysis. However, all the commercially available data recording software for the low-cost thermal camera do not provide users with a function to extract the raw data, that is a sequence of temperature metrics. Available tools produce colour-mapped images which are already converted from raw data. As their own temperature-to-colour mapping formula is not provided (histogram equalization is often applied), it is hard to extract actual temperature values from the colour images. Hence, there is a need to design a new recording tool that can provide a time series of 2D temperature information.

To capture the temporal 2D thermal information, we developed software using FLIR One SDK for Android¹¹, which allows to access to data produced by the thermal imaging controller (in the device) and extract the raw thermal sequences. Upscaling and bicubic interpolation were applied to the original 160x120 thermal matrix to double the size of the matrix and enhance its quality, as shown in Figure 3.3a. We also implemented a function to assist with changing the emissivity coefficient which requires adjustment along with a type of targeted material surface.

¹¹<http://developer.flir.com/flir-one-software-development-kit/>

For all the data collection studies carried out in this thesis, the emissivity was set to 0.98 for the human skin (Steketee, 1973).

Figure 3.3b shows a data frame protocol for mobile thermal imaging as there is no standard protocol for sharing two-dimensional thermal data sequences in the research community. Note that every low cost, mobile thermal imaging device available in early 2018 uses the MPEG protocol which is one of the RGB video/audio compression protocols. The compressed outcomes are already quantised colour-mapped images or videos, not thermal sequences. Therefore, we developed the data frame protocol for storing the raw thermal data sequences.

Temperature sequences stored within the proposed data frame protocol can be easily loaded on Matlab or Python – example codes are available at the author’s GitHub repository¹². Each frame is stored in a 16bit Boolean binary array, instead of 8bit data, to avoid degrading the thermal sensitivity. As discussed in Section 2.1, the low-cost thermal imaging devices (such as FLIR One and Seek Thermal, see Table 2.1) do not produce data at a consistent sampling rate (i.e. unsteady temporal resolution due to their dependency on a mobile operating system) For this reason, time stamps are included in the end of each frame (*time info* in Figure 3.3b).

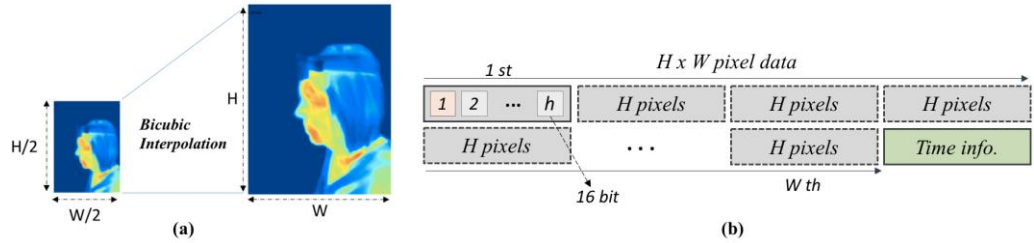


Figure 3.3. (a) Thermal image with bicubic interpolation and scaling, (b) the proposed data frame protocol.

3.3. Overview of Experimental Protocols and Datasets

This section describes key experimental protocols to create bespoke datasets that allow for specifically addressing our research questions, comparing our proposed methods with state-of-the-art approaches. Inspired by the literature, we design two groups of experimental protocols along with task type: i) without and ii) with stress-induction tasks. The first group of protocols is

¹² <https://github.com/deepneuroscience/Thermal-Video-Read>

to collect data from participants in various activities and environments with the main purpose of testing signal processing and computational methods which this thesis proposes and builds for automatic tracking of ROIs and physiological signatures in Stages I and II. Breathing exercise, standing, and sitting (for having a conversation, surfing the internet, reading articles, etc), walking indoors and outdoors are chosen as main activities.

The second group of protocols aims to collect data from participants who specifically attend stressful tasks to ensure capturing of stress-induced physiological patterns in Stage II and to test computational methods for stress recognition in Stage III. As main tasks, Stroop and mental arithmetic tests are used. Both task types were chosen based upon studies already tested in the literature for systematically validating our methods and comparing them with the existing state-of-the-arts in highly constrained situations. However, our protocols were also designed to be conducted in unconstrained situations so as to see how they could reliably work in real-world like situations where environmental temperature changes and natural movements of participants are allowed.

Without stress-induction tasks (Stages I and II - Chapters 4, 5, 6)

In more detail, the experimental protocols without stress-induction tasks were designed with different scenarios along with higher complexity reflecting the thermal dynamic ranges and motion artefacts. These are referred to as:

- (i) controlled breathing exercise with **constrained** head motions in multiple indoor and outdoor places with different ambient temperatures,
- (ii) indoor sedentary activity with **constrained** head motions and room temperatures,
- (iii) indoor sedentary activity with **unconstrained** head motions and room temperatures,
- (iv) indoor and outdoor physical activity in **fully unconstrained, mobile** contexts and with variable thermal range scenes.

The aim of the first and second experiment protocols (i, ii) are for systematic evaluations of our proposed methods. As in Gastel *et al.* (2016), the first protocol uses guided breathing patterns and constrains the person's movement. Following Veltman & Vos (2005), the second protocol uses a chinrest to minimise head motions of participants and controls room temperatures.

Both the third and fourth experimental protocols (iii, iv) are to evaluate our methods under natural, real-world like conditions: the third targets sedentary activities (such as sitting and reading articles as in Abdelrahman *et al.*, 2017) without constraining movement and with low-levels of room ambient temperature variances. To simulate real-world like situations, we use open places (e.g. real office) where strangers are present and natural conversation is allowed. The last protocol targets physical activity in fully mobile contexts with varying ambient temperature. This includes walking through indoor corridors, stepping and strolling in a park area.

With stress-induction tasks (Stages II and III - Chapters 6, 7, 8 and 9)

The experimental protocols with stress-induction tasks are to build datasets which help to validate the proposed approaches under stressful conditions. These are:

- (i) stress-induction desk activity tasks with **constrained** head motions and room temperatures,
- (ii) stress-induction desk activity tasks with **unconstrained** head motions and room temperatures.

The first protocol (i) is mainly for exploring cortical thermal signatures (in Chapter 7). Whilst we have found physiological and neuroscientific evidence demonstrating a possibility that cortical thermal signatures could respond to mental stressors, this has been unexplored in the literature; hence a very constrained protocol is required. On the other hand, the second protocol (ii) is to impose no constraints to participants, allowing their unconstrained movement (e.g. head motion) during the sedentary activities. In addition, room temperatures are also not controlled similarly to the previous group of experimental protocols.

To be specific, during both types of stress induction study protocols, we collect physiological data and subjective self-reports in association with mental stress levels due to mental workload (Lazarus, 1993; Everly Jr & Lating, 2012; Alberdi *et al.*, 2016). We have chosen two cognitive-load induction tasks to build our study protocols - *the Stroop Colour Word test* (Stroop, 1935) and *the Mathematical Serial Subtraction test* (Soufer *et al.*, 1998). These tests are mainly selected as they have been shown in the literature on mental stress induction studies in psychology, neuroscience and affective computing with the aim to induce mental stress by increasing cognitive load (Åkerstedt *et al.*, 1983; Dedovic *et al.*, 2009; Hong *et al.*, 2012; McDuff *et al.*, 2016). They have also been used in other thermal imaging studies (Puri *et al.*, 2005; Engert

et al., 2014). Each task needs to be divided into two sub-tasks with varying difficulty levels in order to elicit different stress levels, for instance, easy and hard: Se – Stroop easy, Sh – Stroop hard, Me – Math easy, Mh- Math hard. Each sub-task is counterbalanced in a Latin squared design.

Although it has been shown that the Stroop and mental arithmetic tests lead to cognitive overload (Åkerstedt *et al.*, 1983; Taylor *et al.*, 1997; Soufer *et al.*, 1998; Dedovic *et al.*, 2009; Alberdi *et al.*, 2016), they are limited in the amount of stress they induce due to the lack of psychosocial stressors or other stressors (Dedovic *et al.*, 2009; Setz *et al.*, 2010). Hence, following the literature (Åkerstedt *et al.*, 1983; Genno *et al.*, 1997; Dedovic *et al.*, 2009; Setz *et al.*, 2010), we also introduce further stressors: a) social evaluative threats, i.e., close observation and assessment of a person's performance (Dedovic *et al.*, 2009; Setz *et al.*, 2010), b) time pressure, for example, 1.5 second limitation for each Stroop question (Åkerstedt *et al.*, 1983), and c) loud sound feedback, in particular, an unpleasant sound for wrong answers (Genno *et al.*, 1997).

To collect subjective self-reports of mental stress levels, we commonly use a 10-cm Visual Analogue Scale (VAS), which allows participants to answer on an analogue basis (continuous) to avoid non-parametric properties (Bijur *et al.*, 2001; Lesage *et al.*, 2012). The main question asked is “*How much did you feel mentally stressed?*” (ranging from 0, not at all, to 10, very much). Only one VAS straight line is used in the stress induction data collection studies in this thesis to self-report his/her perceived stress levels across all tasks and sessions. This is to help participants easily compare stress scores they report between sessions as shown in Figure 3.4 where a 10cm VAS based questionnaire is visualised with examples of labels along with each session type (R₁, R₂: Rest from Session 1 and 2, Se: Stroop easy, Sh: Stroop hard, Me: Math easy, Mh: Math hard). This approach combines a numerical approach to self-reporting with a ranking one, as ranking is generally more reliable than simple quantisation of a subjective state (Bang, 2009; Atkinson *et al.*, 2016; Yannakakis *et al.*, 2017). Examples of labels (R₁, R₂: Rest from Session 1 and 2, Se: Stroop easy, Sh: Stroop hard, Me: Math easy, Mh: Math hard) in Figure 3.4 have been added to the figure by the researcher to clarify their reference to each task. The entire questionnaires used in this thesis is provided in Appendix H.

Q) “How much did you feel mentally stressed?”

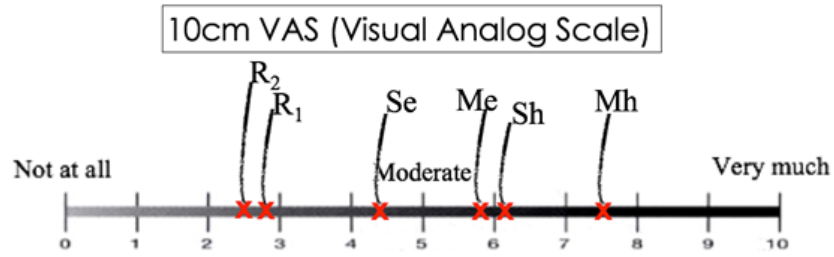


Figure 3.4. 10cm VAS based questionnaire with examples of labels along with each session type (R₁, R₂: Rest from Session 1 and 2, Se: Stroop easy, Sh: Stroop hard, Me: Math easy, Mh: Math hard). The red marks (x) represent examples of self-reported scores of one participant over the different tasks. The task labels have been added by the researchers for the purpose of this figure.

Summary of collected datasets

Table 3.2 summarises our collected datasets using the experimental protocols described above. This is to provide a quick glance of our incrementally built labelled datasets of thermal images in both constrained and everyday ubiquitous settings. Each detail is provided in each corresponding chapter in terms of task flows, information of participants, the number of thermal cameras used, and other commercialised physiological sensors used for the purpose of technically evaluating proposed methods. In addition, we provide detail of how we build a richer set of physiological data sets (multiple types of physiological data were collected from additional sensors for building richer datasets even if they are not exploited in this thesis). Prior to all data collection conducted in this thesis, participants were asked to read the information sheet and sign the informed consent form attached in Appendix I. All the experimental protocols described above were approved by the Ethics Committee of University College London Interaction Centre (ID Numbers: STAFF/1011/005 and UCLIC/1617/003). Participants were informed that they could stop the study at any time if they felt uncomfortable.

Table 3.2. An overview of an incrementally built, labelled datasets from conducted data collection studies in constrained systematic and unconstrained settings

<i>Dataset Name</i>	<i>Chapter</i>	<i>Constrained Systematic Experiments</i>	<i>Unconstrained Experiments</i>
<i>ROITracking</i>	4 & 5	-	indoor sedentary activity with unconstrained head motions and room temperatures (N=10)
<i>ThermalBreath I</i>	5	controlled breathing exercise with constrained head motions in multiple indoor and outdoor places with different ambient temperatures (N=5)	-
<i>ThermalBreath II</i>	5	-	indoor and outdoor physical activity in fully unconstrained, mobile contexts and with variable thermal range scenes (N=8)
<i>NoseTracking</i>	6	indoor sedentary activity with constrained head motions and room temperatures (N=10)	-
<i>StressNose</i>	6	-	stress-induction desk activity with unconstrained head motions and room temperatures (N=12)
<i>ThermalBrain I</i>	7	-	stress-induction desk activity with unconstrained head motions and room temperatures (N=10)
<i>ThermalBrain II</i>	7	stress-induction desk activity with constrained head motions and room temperatures (N=10)	-
<i>DeepBreath</i>	8	-	stress-induction desk tasks with unconstrained head motions and room temperatures (N=8)
<i>InstantStress</i>	9	-	stress-induction desk tasks with unconstrained head motions and room temperatures (N=17)

3.4. Overview of Evaluation Methods

This thesis thoroughly evaluates our proposed approaches using incrementally built datasets from the experimental protocols described in Section 3.3. We reimplement existing state-of-the-art approaches for comparing performance of our proposed approaches, particularly in Stages I and II (i.e. automatic ROI tracking, physiological signal tracking). We also implemented existing standard metrics to represent such physiological signatures to compare them with our proposed representation and metrics. In the case of physiological signature tracking, reference data from commercial physiological sensors are adopted as a ground truth to statistically evaluate results. Results from the stress induction experiments, such as self-report scores and physiological thermal signatures, are statistically analysed.

To be more specific, in Stage I, we implement existing ROI tracking methods which are most widely used in the computing literature such as the Median Flow (Kalal et al., 2010). Two facial areas, periorbital and perioral areas, are chosen as ROIs to track as both areas have been widely explored in the literature (e.g. Pavlidis *et al.* 2002). The number of successfully tracked frames is used as a metric to assess the performance. It is programmed to store a movie with the tracking result from every frame of thermal image so as to facilitate visual inspections. In the case of trackers based on “*Good Features to Track*” (Shi & Tomasi, 1994) (a majority of state-of-the-art algorithms are built on this, such as Median Flow), the tracking evaluation program is programmed to finish and report the metric value when all feature points are lost. As the proposed signal processing technique is a backbone of computational methods built in Stages II and III, as well as being thoroughly tested in terms of motion artefacts and variable environmental temperatures, this stage uses only the dataset collected from the indoor sedentary activity with unconstrained head motions and room temperatures and track the two ROIs. Lastly, in order to test the general capability of the method in other material temperatures than the human skin, we additionally conduct a material classification task by collecting numerous outdoor materials. For the evaluation, we implement a state-of-the-art deep learning algorithm (Spatial Transformer-based convolutional neural network, Jaderberg *et al.*, 2015) and compare the classification performance between the proposed processing-enabled method and the original algorithm.

In Stage II, we additionally implement an existing state-of-the-art ROI tracking method, Sparse Representation (Mei & Ling, 2011) used in the state-of-the-art breathing tracking method proposed by Pereira *et al.* (2015). As we conduct in the first stage, we use the tracking success rate and evaluate each performance statistically using significance tests (e.g. Friedman’s analysis

of variance, ANOVA). Furthermore, we implement the dominantly used signal representation technique, which is to represent two-dimensional data within a ROI box via averaging (Murthy & Pavlidis, 2006; Fei & Pavlidis, 2010; Abbas *et al.*, 2011; Lewis *et al.*, 2011; Gane *et al.*, 2011; Pavlidis *et al.*, 2012; Pereira *et al.*, 2015; Abdelrahman *et al.*, 2017), and metrics to quantify physiological thermal signatures such as temperature difference or respiratory rate. They are used as benchmarks to compare the quality of physiological signatures tracked by the proposed methods. At the same time, we evaluate the results with reference signals from existing physiological sensing devices (e.g. chest respiration belt) with statistical tools such as the Bland-Altman plot and the root mean square error (RMSE). In the case of data collection studies using stress-induction studies, we statistically compare between self-report scores and physiological data, or between such data across each condition (e.g. stressful or not stressful) by conducting significance tests. The within-group design is commonly used in this thesis.

In Stage III, we implement existing machine learning classification models, such as k-Nearest Neighbour, a single hidden-layer neural network and convolutional neural network. We test such models with the widely-used gross statistical features from physiological signatures as well as new models based on feature representation techniques this thesis proposes. A set of subjective ratings on questionnaires collected from stress induction-tasks are used to label data and set ground truths. Finally, a Leave-One-Subject-Out (LOSO) cross-validation technique is used to test their ability to generalise towards different individuals.

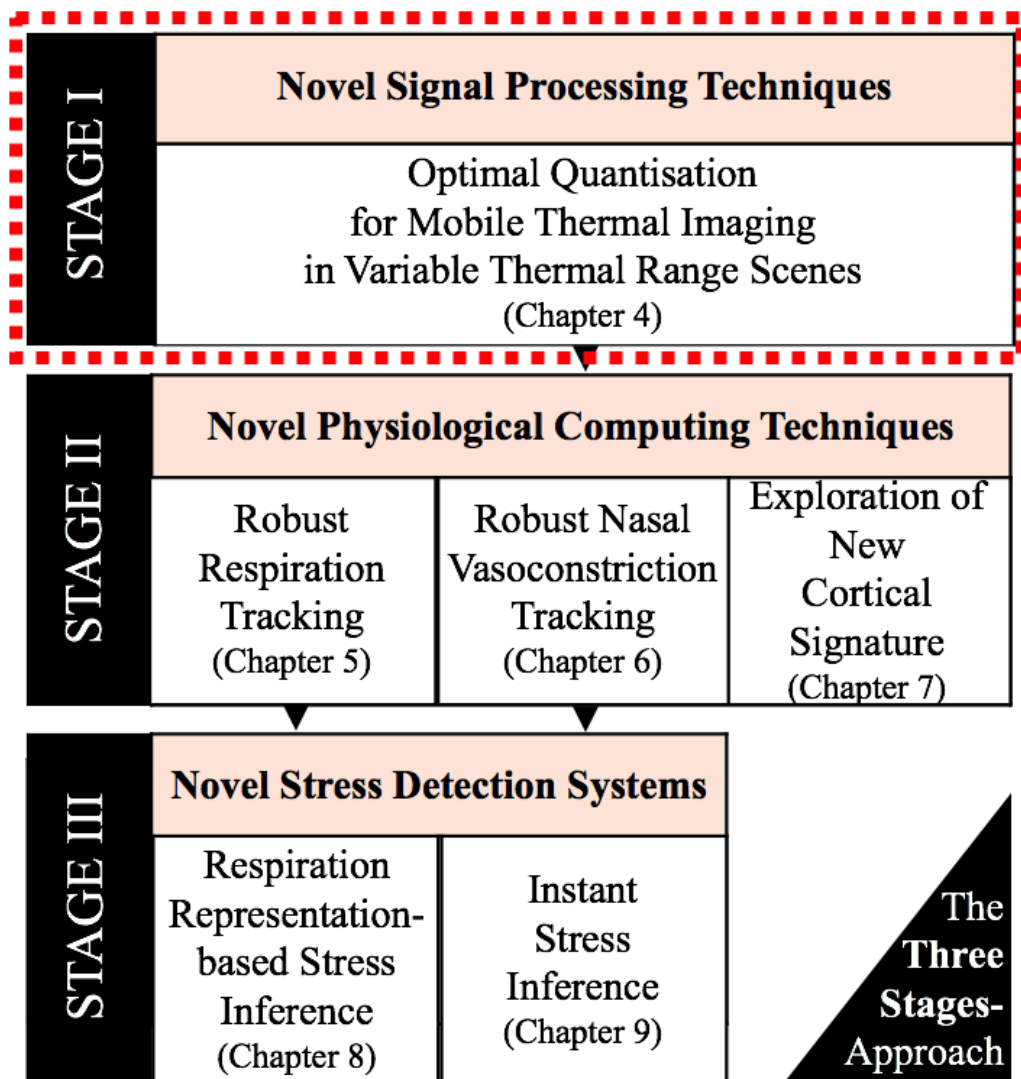
3.5. Chapter Summary

This chapter is dedicated to the establishment of our research methodology to bringing thermal imaging from constrained, systematic environments into unconstrained, mobile, real world situations. The key approach is stacked through the three stage strategy. Earlier stages are prerequisites to enable every next stage towards answering the key research question of this thesis: *“Can mobile, low-cost thermography be used as multiple physiological measures for automatic recognition of a person’s mental stress in HCI settings?”*. Basic tools have been built to support the use of low-cost thermal cameras as a main physiological sensing channel. Experimental protocols have been designed to incrementally produce datasets considering the problem which each proposed method is to solve along with each stage and chapter. We have also discussed evaluation methods to appropriately assess the performances of our proposed approaches upon

the built datasets. This chapter is followed by Chapter 4 which stands on the first stage. In the following chapter, novel signal processing techniques are proposed and evaluated, with the aim of enabling the use of thermal imaging in variable thermal range scenes.

STAGE I

Novel Signal Processing Techniques
for Mobile Thermal Imaging



Chapter 4

Optimal Quantisation: A Preprocessing method for Mobile Thermography in Variable Thermal Range Scenes

As discussed in Chapter 2, to bring thermal imaging from constrained, systematic settings into unconstrained, ubiquitous applications, we ought to confront the key challenge: variable environmental temperature. This negatively affects graphical and morphological structures of skins or materials on thermal images, leading to difficulties in automatically tracking ROIs, which is a prerequisite to reliably, continuously and automatically extract physiological signatures. To address this key issue, this chapter aims to build a new pre-processing method that can support the use of thermal imaging in high-thermal dynamic range scenes.

Section 4.1 discusses the background and our research question. Section 4.2 proposes the novel signal processing method and the method-enabled ROI tracking system. Sections 4.3 and 4.4 describe datasets and evaluation methods, which are followed by Section 4.5 where performance of tracking human skin ROIs is presented. Section 4.6 further explores how the proposed method can be generalised to other material temperatures. The chapter ends with a discussion and a summary in Section 4.7 and 4.8.

4.1. Background and Research Questions

Whilst in the majority of earlier work researchers have carefully and manually tracked a person's ROI on recorded thermal image sequences to observe temperature variances induced by certain affective states, only a few studies have applied automatic ROI (or motion) tracking algorithms -

which have been developed for RGB video analysis - to thermal sequences (Garbey et al., 2007; Sun et al., 2005; Ebisch et al., 2012; Manini et al., 2013; Shastri et al., 2012; Tsiamyrtzis et al., 2007; Zhou et al., 2009). Despite documented motion artefacts leading to less reliable physiological signal acquisition from thermal imaging, recent greater improvements made in vision-based ROI tracking (Kalal *et al.*, 2010, 2012; Mei & Ling, 2011) are also promising for thermal imaging, possibly helping to handle the issue. However, beyond controlled laboratory settings, another key computational challenge is environmental thermal dynamics which affect morphological, graphical properties of thermograms.

This is a similar issue to illumination-related issues of RGB vision imaging: for example, beyond typical situations with moderate light levels, an RGB camera used for physiological monitoring (Kumar *et al.*, 2015) or material recognition sensitivity (Liu *et al.*, 2010; Aittala *et al.*, 2015) do not work properly. The environmental temperature issue has never been addressed in the literature up to this point as their high-cost, immobile thermal cameras have limited their use to highly constrained lab settings where ambient temperature is controlled (Garbey et al., 2007; Sun et al., 2005; Ebisch et al., 2012; Manini et al., 2013; Shastri et al., 2012; Tsiamyrtzis et al., 2007; Zhou et al., 2009). Thus, this needs to be tackled to make thermal imaging less thermal illumination-dependent.

In detail, despite our homoeothermic characteristics, temperature distribution on the cutaneous skin of the human face changes with the ambient temperature. Therefore, the changes in thermal patterns could lead to difficulties in identifying each facial area on thermal images. Figure 4.1 for instance, shows thermal images of a person being exposed to ambient air while walking outside during the summer. As can be seen, as the person's face cools down, information about facial morphology becomes lost. However, this is a direct result of the way the image was created. The thermal camera measures temperature directly, and this must be quantised or converted into a digital signal with a fixed number of bits. The quantisation process must cover where the bulk of the signal lies.

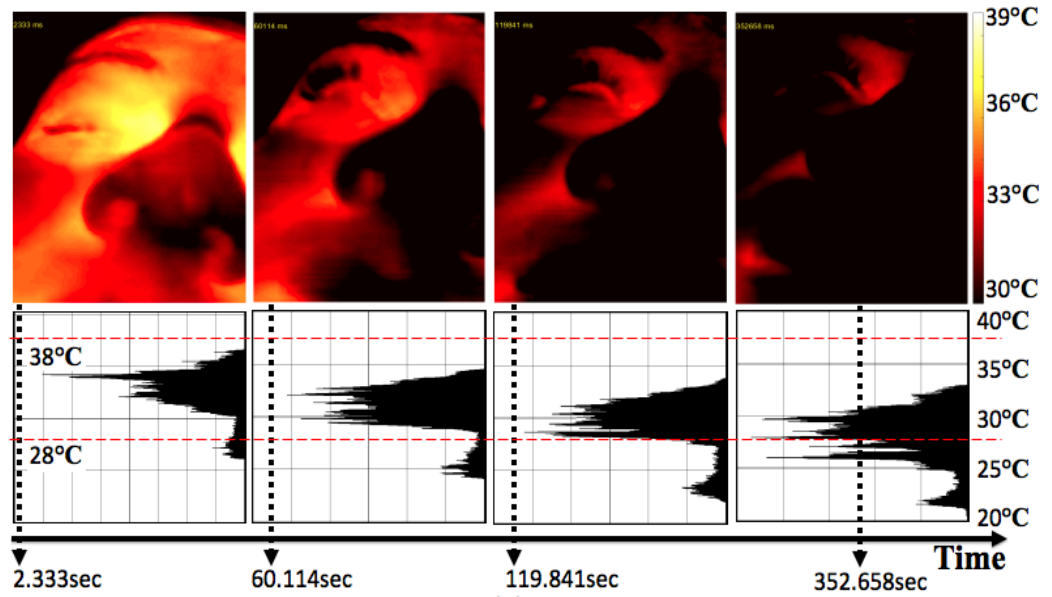


Figure 4.1. Key challenge in the ROI tracking through thermal imaging in high thermal dynamic range scenes: fixed thermal range of interest is not suitable in preserving the morphological facial shape within varying ambient temperature: [top] examples of thermogram shots collected from a person walking outdoor (for 6 minutes), [bottom] temperature histograms.

The most common approach is to use linear quantisation with a selected temperature range of interest, which is traditionally fixed from the first thermogram frame (e.g. 28°C to 38°C in Fei & Pavlidis, 2010). However, this is unable to adapt to the dynamic situation here in which the temperature falls below 28°C in many parts of the image. One way to address this would be to increase the thermal range of values in colour-mapping to ensure all salient features are captured. However, because the resolution of the digital signal is fixed, this causes the resolution to go down. Because the range of temperature distribution over the facial area is relatively narrow, this would reduce contrast and the ability to detect subtle signals. These factors directly influence automatic ROI tracking performance.

This leads to the research questions: *How can we address the effect of variable thermal range scenes, which is one of the key challenges in automatic tracking of ROIs (Regions of Interest) in real world, mobile situations?* To answer this, we focus on building a new solution to colour mapping, which can continuously adapt to consider high dynamical thermal changes in order to improve segmentation, classification, and in turn ROI tracking. Furthermore, the last part of this chapter is dedicated to a demonstration of the general capability of the proposed new solution to other material temperatures in other HCI applications and material classifications, in

order to answer the second question: *How can we generalise the proposed approach to other material temperatures than human skin temperatures?*

4.2. Proposed Computational Methods

In this section, we propose a novel signal processing technique to help compensate for the negative effect of variable thermal range scenes (Section 4.2.1) and then we describe how the technique can support a ROI tracking algorithm (Section 4.2.2).

4.2.1. Optimal Quantisation Algorithm

In thermal image processing, Quantisation is the process of translating from a continuous temperature value to its digital colour-mapped equivalent. Suppose the temperature lies in the range $[T_0, T_{k-1}]$ and the colour pixel which represents it lies in the range $[u_0, u_{k-1}]$. We define $u = \Gamma(T)$ as the mapping between the two. For the Quantisation in time-varying thermal dynamic range scenes, our idea is to adaptively quantise the thermal distribution sequences by finding a thermal range of interest that contains the whole facial temperature distributions for every single frame.

As a first step, we use a statistical extreme value removal process to reduce unexpected noises (e.g. sunlight projected on a person's glasses) and points of extreme temperature, which may be produced by errors of mobile thermal imaging in calculating temperature due to fixed emissivity or lens conditions (e.g. misted lens). By removing thermal signals beyond 1.96 standard deviations, we can set an initial candidate for the thermal range $[T'_{\min}, T'_{\max}]$:

$$T'_{\min} = \bar{c} - 1.96 \frac{\sigma_c}{\sqrt{n \cdot m}}, \quad T'_{\max} = \bar{c} + 1.96 \frac{\sigma_c}{\sqrt{n \cdot m}} \quad (4.1)$$

where \bar{c} is the sample mean of $c(x)$ which is the one-dimensional temperature distribution ($\forall x \in 1, n \cdot m$), $n \cdot m$ is the spatial resolution of collected thermal distribution matrixes (Figure 4.1), and σ_c is the standard deviation of $c(x)$.

Now for the final selection the interest range, we assume there are two qualitatively different elements in each frame: the person's face and the background. Therefore, we adopt Ridler and Calvard's concept of optimal threshold selection (Ridler & Calvard, 1978). This

method finds the threshold value that best separates objects from the background by iteratively analysing the colour histogram. This can help to search thermal values that distinguish the human skin from non-skin areas and the time-varying temperature histograms exhibiting various dynamic ranges. Furthermore, it is known this method is robust, even in the presence of non-bimodal histograms which make it harder to find an optimal boundary of temperature ranges of interest. Our Optimal Quantisation technique, which is named after the concept, is finalized with the iterative computation of an optimal threshold value T_{opt} below:

$$T_{opt}(0) = T'_{\min} \quad (4.2)$$

$$T_{opt}(t+1) = \frac{\mu_1(t) + \mu_2(t)}{2} \quad (4.3)$$

where $\mu_1(t)$ and $\mu_2(t)$ are the mean values when $c(x) \leq T_{opt}(t)$, $c(x) > T_{opt}(t)$, respectively. While the method presented in Ridler & Calvard (1978) requires the four corners of an image to contain background pixels, our method does not require this condition because Equation (4.2) is initialised with a value from Equation (4.1). The process is iterated until $T_{opt}(p) - T_{opt}(p-1) \approx 0$ is satisfied and the temperature range of interest is then chosen to be

$$T_0 = T_{opt}(p), \quad T_{k-1} = T'_{\max} . \quad (4.4)$$

When the average temperature over the background, which includes hair and air, is lower than that of human cutaneous skin, only the lower bound is of interest in determining the optimal range. When the average temperature of the background is above that of human cutaneous skin, the upper bound can be used instead.

4.2.2. Optimal Quantisation-powered Automatic ROI Tracking

The aim of this section is to present how the proposed quantisation method can enable existing automatic ROI tracking methods (which are for RGB images) to perform on thermal images. A series of computations following the recording phase contain: i) the Optimal Quantisation, ii) ROI Tracking, iii) extraction of original 2D thermal matrices from the tracked ROI, and iv) a post processing, as summarised in Figure 4.2. We used the Median Flow algorithm (Kalal et al., 2010), one of the state-of-the-arts, for tracking point features on colour converted images created by the quantisation technique (Figure 4.2b, 4.2c). Details of the tracking process are described in Chapter 5 (Section 5.2.1) where we explore how to further improve the existing tracking

algorithm. This process results in capacitating the extraction of a two-dimensional apparent temperature sequence from the tracked ROI (Figure 4.2d). Although explorations of thermal variances from the ROI are beyond the scope of this chapter, please note that this can be followed by a post-processing step which is necessary. For example, sparse signal restoration that estimates null values between sparse elements (Figure 4.2e) which addresses unsteady temporal frame rates of mobile thermal imaging, as discussed in Section 2.1. In addition, thermal voxel integration and outlier removals which are described in the following chapters.

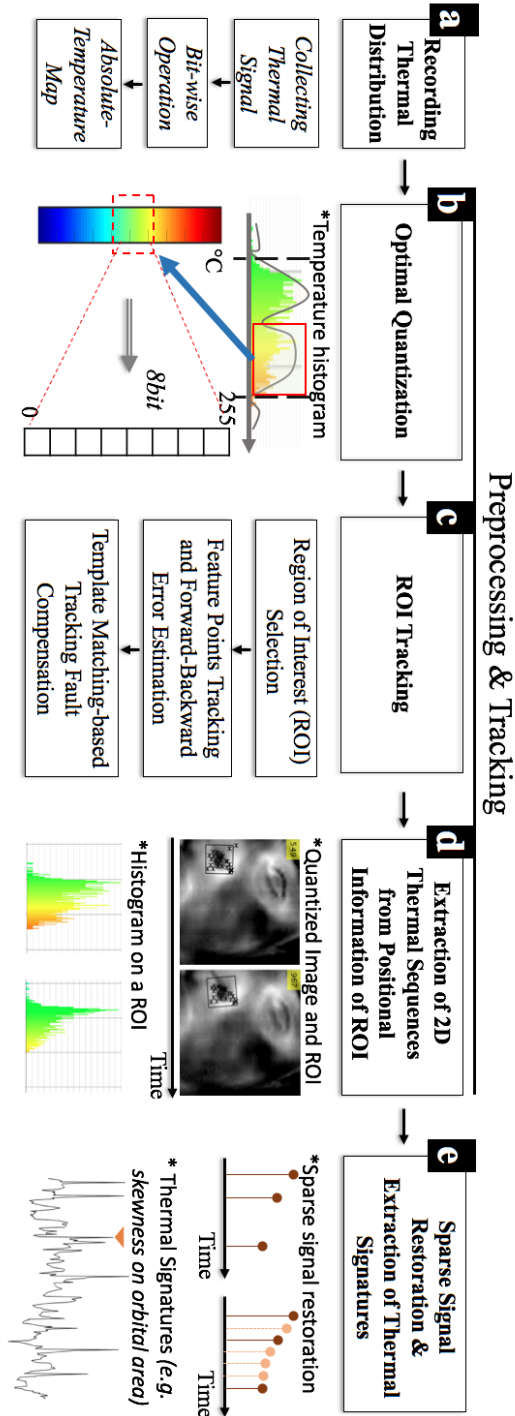


Figure 4.2. The proposed optimal quantisation and a processing flow for automatic tracking of a region of interest (ROI) and the extractions of thermal variances from the ROI.

4.3. Experimental Protocol and Dataset (*ROITracking*)

This section describes the experimental protocol and the collected dataset (named *ROITracking*). This aims to collect facial thermal images from participants taking part in indoor sedentary activities with **unconstrained** room temperatures and head motions. This is to mainly test the Optimal Quantisation-powered automatic ROI tracking performance on human skin areas. As in Abdelrahman *et al.* (2017), desk activities such as sitting and reading articles are used. In detail, 10 healthy adults (6 female) (aged 24-31 years, $M=28.4$, $SD=2.17$) from a variety of ethnical backgrounds (skin colour: from pale white to black) were recruited from the university subject pool. The pool includes people from outside the university. Participants went through a sequence of tasks consisting of three phases lasting 2 minutes each: i) *sitting and conversation*, ii) *reading a news article on the screen* and iii) *surfing the internet with the keyboard and mouse*.

As described in Figure 4.3, we used the low-cost thermal camera described in Section 3.2 (Figure 3.2) and installed it near each participant's face using a shoulder rig. The distance between a person's face and the device ranged from 35cm to 50cm to account for the low spatial resolution (160x120) of the camera. Each participant was asked to take off glasses and their cap (if any) to clearly capture thermal variations on their facial cutaneous skin including the periorbital and perioral areas. The data collection study was conducted in an office room to simulate everyday desk activities. Although a chest-belt type respiration sensor was used in this data collection, the data is not used in this chapter (this is to evaluate performance of respiration tracking methods which are discussed in the following chapter).

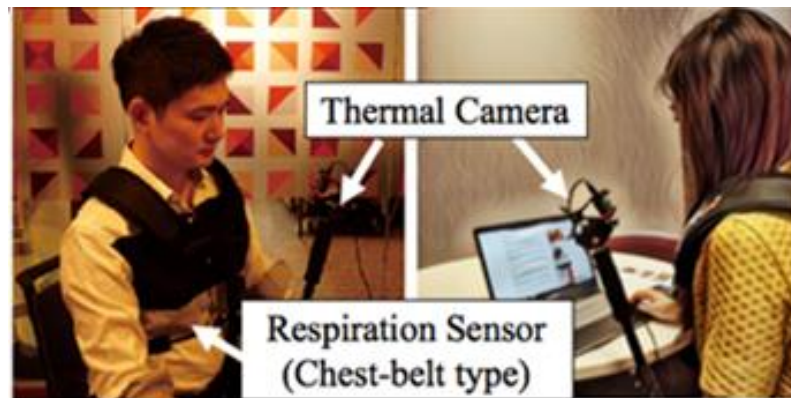


Figure 4.3. Experimental setup: a low-cost thermal camera was installed near each participant's face using a shoulder rig. An office room was used to simulate everyday desk activities.

Participants were not asked to constrain their movements. Additionally, a chest-belt type respiration sensor was used. The data is not used in this chapter but used as a reference signal in the following chapter where we evaluate respiration tracking algorithm performance.

To simulate natural behaviours during in-situ tasks, behaviour such as turning of the head during a conversation, body shaking due to laughing and any other natural movements were not controlled or restricted in any way. As shown in Figure 4.4, this resulted in collecting thermal image recordings including a variety of movement situations, such as head rotations due to people walking behind them with temporary disappearance of certain ROIs, such as the periorbital area or nostrils, from the thermal camera view. The change in a participant's position from the first phase (sitting and conversation) to the second phase (reading an article) ensured sudden head movements; in particular, participants with low vision, approached the screen as shown in Figure 4.4 (top-left and top-middle).

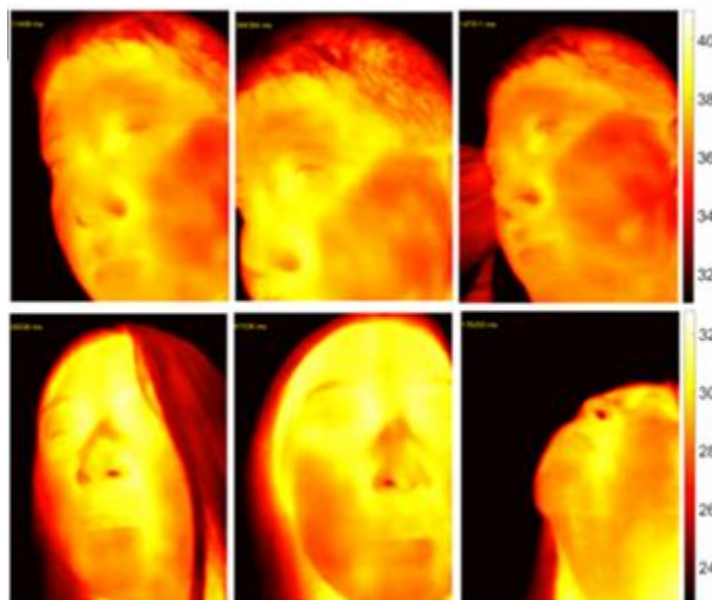


Figure 4.4. Examples of collected facial thermal images of participants. Not constraining participants' movements helped to capture their very natural motions such as lowering, turning, and rotating their heads.

All sessions were run during the summer. To ensure ambient temperature variations, we changed (lowered or increased) air conditioning options during the course of each session. The experiment resulted in circa 60 minutes (10 participants x about 6 minutes) of thermal video recording of natural movements, spontaneous breathing patterns in sedentary contexts and changes in ambient temperature. The total number of frames of thermal recordings for each person ranged from 2558 to 3506 frames (about 6 minutes). Note that the sampling rate of the low cost thermal camera is not fixed, leading to the different range of collected frames.

4.4. Evaluation Method

To evaluate performance of the proposed Optimal Quantisation method in the skin ROI tracking task, we implemented two ROI tracking algorithms. Firstly, Mode-Seeking (also called mean-shift) algorithm (Cheng, 1995) was selected since it is known as a simple and effective traditional method. Secondly, Kalal *et al.*'s Median Flow (2010) was implemented as, given its high tracking success rates, it has been used in latest works employing thermal imaging (Abouelenien *et al.*, 2017) and also used as the base of the proposed Optimal Quantisation-powered method in Section 4.2.2. Since the employed methods do not handle the dynamic quantisation issue, the thermal range of interest to $\pm 5^{\circ}\text{C}$ from the average temperature over a person's whole face in the first thermogram frame was used for static quantisation as in Fei & Pavlidis (2010) (i.e. the range is $[28^{\circ}\text{C}, 38^{\circ}\text{C}]$).

In this evaluation process, periorbital and perioral areas were chosen as main ROIs because they have been the main facial parts having emerged from the literature which explores affect-related thermal signatures (Pavlidis *et al.*, 2001; Veltman & Vos, 2005; Gane *et al.*, 2011; Pavlidis *et al.*, 2012; Engert *et al.*, 2014). To evaluate performances of the proposed and implemented algorithms, we used the number of successfully tracked frames as a main evaluation metric (and its rate over the total frame number) and investigated effects of tracking methods on the metric using the Friedman test.

4.5. Results

Table 4.1 shows the overall results from the collected thermal sequences. We compared each successful result (the number of ROI frames which were successfully tracked) in tracking the two local facial regions. As a result, the tracking performance of the Optimal Quantisation powered Median Flow over-performed the others (Periorbital area: Mode-Seeking $M=49.98\%$, $SD=29.96(\%)$; Median Flow $M=55.25\%$, $SD=26.09(\%)$; **Proposed** $M=99.43\%$, $SD=1.81(\%)$; Perioral area: Mode-Seeking $M=47.23\%$, $SD=33.21(\%)$; Median Flow $M=34.81\%$, $SD=40.15(\%)$; **Proposed** $M=83.11\%$, $SD=25.39(\%)$). To double check the results, we visually inspected each recorded thermal video with the tracker (bounding box).

Furthermore, we tested statistical significance of tracking results from each ROI. For this we used Friedman test (with an alpha value of 0.05) which is the non-parametric alternative to the ANOVA with repeated measures as not all the data for each level of the within subject factor

were normally distributed (e.g. Median Flow with Optimal Quantisation for the periorbital area: $p=0.019$ from the Shapiro-Wilk normality test). The number of successfully tracked frames showed significant effects of the algorithm type (Periorbital area: $\chi^2(2)=14.8$, $p=0.001$, Perioral area: $\chi^2(2)=7.824$, $p=0.02$). For the post hoc analysis of paired conditions, we used the nonparametric Wilcoxon signed ranks test in which the alpha value was adjusted ($\alpha=0.017 \approx 0.05/k$ where $k=3$, the number of pair-wise comparisons). The results showed that the Optimal Quantisation powered algorithm significantly outperformed the other two for the periorbital area (mode-seeking vs. proposed: $Z=-2.521$, $p=0.012$; median flow vs. proposed: $Z=-2.666$, $p=0.008$) and the median flow for the perioral area (median flow vs. proposed: $Z=-2.521$, $p=0.012$; in the case of mode-seeking vs. proposed: $Z=-2.1$, $p=0.036$). The result is very interesting as it indicates our Optimal Quantisation itself can improve tracking performance of existing algorithms in situations where environmental temperature and participants' mobility are not controlled.

Table 4.1. Tracking results: the number of successfully tracked frames – Mode seeking (Cheng, 1995) vs. Median Flow (Kalal *et al.*, 2010) vs. Median Flow with the proposed Optimal Quantisation (bold: perfect).

Participants	Total frame	ROI: Periorbital area			ROI: Perioral area		
		Mode-seeking	Median Flow	Median Flow with Optimal Quantisation	Mode-seeking	Median Flow	Median Flow with Optimal Quantisation
P1	3195	509	737	3195	360	638	3195
		15.93%	23.07%	100%	11.27%	19.97%	100%
P2	3282	3282	731	3282	3282	81	3282
		100%	22.27%	100%	100%	2.47%	100%
P3	3183	532	1545	3183	159	20	1281
		16.71%	48.54%	100%	5.00%	0.63%	40.25%
P4	3404	1576	1747	3404	1753	3404	3404
		46.30%	51.32%	100%	51.50%	100%	100%
P5	3152	1185	1237	3152	966	523	2924
		37.60%	39.24%	100%	30.65%	16.59%	92.77%
P6	2558	982	1551	2411	874	297	2411
		38.39%	60.63%	94.25%	34.17%	11.61%	94.25%
P7	3506	1318	2955	3506	1050	2593	3506
		37.59%	84.28%	100%	29.95%	73.96%	100%
P8	3368	2238	2717	3368	2323	169	2208
		66.45%	80.67%	100%	68.97%	5.02%	65.56%
P9	3011	1230	1278	3011	1228	539	1152
		40.85%	42.44%	100%	40.78%	17.90%	38.26%
P10	3238	3238	3238	3238	3238	3238	3238
		100%	100%	100%	100%	100%	100%

4.6. Generalisation to Other Material Temperatures

Beyond the human skin temperature which this thesis mainly investigates, this section explores the general capability of the proposed Optimal Quantisation in analysing material temperatures which can be found from highly variable temperature environments. As explored in the computing literature (Larson *et al.*, 2011; Kurz, 2014; Sahami Shirazi *et al.*, 2014), thermal imaging with high-cost, heavy thermographic systems has also been employed in other interesting HCI applications which have mostly targeted office environments. With the use of low-cost thermal cameras helping improve its accessibility to the general public, we expect that the Optimal Quantisation can help to tackle the boundary using thermal imaging and bring this to more wild, ubiquitous applications.

Here, for the first time, we bring mobile thermal imaging outdoors, specifically aiming to create the ability to automatically recognise materials for which a variety of sensing technology have been investigated in HCI (Bell *et al.*, 2015; Sato *et al.*, 2015; Yeo *et al.*, 2016; Wang *et al.*, 2016). This is to support context and material awareness which are of critical importance in many situations for both humans and ubiquitous technologies (Schmidt *et al.*, 1999; Dourish, 2004; Dey & Häkkinen, 2008). For doing this, we use the proposed Optimal Quantisation which is to compensate for high thermal dynamic range scenes, together with one of the latest machine learning techniques. In the following sub-section, we introduce our new mobile thermal imaging-based proximate material type recogniser that can work in a wide range of environmental illumination, temperature, and viewpoints.

4.6.1. Method: Deep Thermal Imaging

Deep Thermal Imaging is a novel approach for sensing and recognising the material of proximate objects. Rather than focusing on the overall material temperature, we propose to focus on the quantisation process to capture the characteristic spatial thermal pattern of the compounds or elements forming a material (Callister & Rethwisch, 2011). Figure 4.5 illustrates our four-step approach: a) physical sensing of proximate surfaces through mobile thermal imaging; b) dynamic range Quantisation to decrease environmental temperature effects and enhance thermal texture; c) deep-learning-based feature learning and training; and d) real-time recognition of materials.

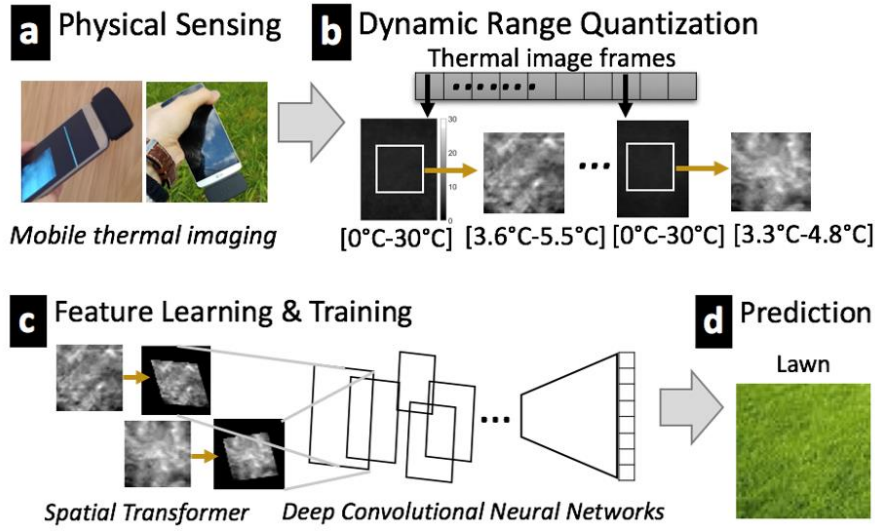


Figure 4.5. The overall process for proximate material type recognition with Deep Thermal Imaging: a) physical sensing via mobile thermal imaging, b) dynamic range quantisation on every thermal image frame, c) feature learning and training process based on spatial transformer and deep convolutional neural networks, d) prediction results.

Physical Sensing: Mobile Thermal Imaging

When attached to a smartphone, a low-cost, mobile thermal camera can be used to continuously sense and record thermal image sequences of different types of surface materials as shown in Figure 4.5a. Its portability means that it is easy to gather images of the same object from many different perspectives (e.g., shape, geometry, illumination) and distances.

Dynamic Range Quantisation

Various factors affect the thermal texture of materials captured by the camera. Each materials compound combinations, surface structure (e.g., roughness) and geometry (e.g., cavities) influence both the amount of radiation emitted from the surface and the emissivity (Vollmer & Klaus-Peter, 2017). Given this, thermal imaging detects the thermal radiation and *estimates* the temperature by using an emissivity value, so-called *apparent temperature* (not *true temperature*). Instead of attempting to measure the true temperature of the material, we aim to capture the thermal radiation differences over material surfaces to discriminate them.

To maximise the differences between materials and compensate for environmental temperature changes, we introduce the *Dynamic Range Quantisation* method, simplified from the *Optimal Quantisation* method proposed in Section 4.2. Quantisation, the process of mapping

input data to output data within a specific range, is crucial for handling thermal image sequences in the wild as discussed earlier. Our method amplifies the *relative* difference between neighbouring temperature elements by taking into account the different amounts of radiation on every single thermal image of a material surface as shown in Figure 4.5b.

First, it crops an $N \times N$ square at the centre of the captured image. In our implementation, we use a 75×75 square given the low spatial resolution (160×120) of the thermal camera. Cropping is important because low-cost systems, designed for mobile thermal imaging, suffer from lens-inducing calibration errors that can lead to extremely low or high temperature at the edge of the image. The second step is to adaptively quantise the thermal measurements using the expression $u = \Gamma(T)$, where $[T_1, T_2]$ is the range of temperature and $[u_1, u_2]$ is the range of a pixel value on the output thermal image. The temperature range is given by

$$T_1(k) = \min_{\forall i, j \in ROI} S_{ij}(k), \quad T_2(k) = \max_{\forall i, j \in ROI} S_{ij}(k) \quad (4.5)$$

where $S_{ij}(k)$ is the element (i, j) of the thermal matrix S with frame number k . The final quantised value $I_{ij}(k)$ on the processed thermal image is:

$$I_{ij}(k) = (u_2 - u_1) \cdot \left(\frac{S_{ij}(k) - T_1(k)}{T_2(k) - T_1(k)} \right). \quad (4.6)$$

This conversion has two effects. Firstly, it produces an output image with more distinct spatial thermal patterns of material as shown in Figure 4.5b, which help to discriminate between the different types. Secondly, these patterns are more invariant to changes in ambient temperature.

Feature Learning Process using Deep CNN

Given the parallel to image processing with RGB images, our feature learning process uses the deep Convolutional Neural Network (CNN) illustrated in Figure 4.5c. Many CNNs have been proposed for RGB-vision analysis (Cimpoi *et al.*, 2015; Jaderberg *et al.*, 2015; Ren *et al.*, 2015; Wang *et al.*, 2016). In this work, we use the Spatial Transformer-based CNN (Jaderberg *et al.*, 2015). This architecture gets its name from the fact that the Spatial Transformer artificially manipulates spatial data information within CNNs. This helps to handle high-level variances of spatial thermal image patterns due to different shooting perspectives and geometries. This technique also performs well on lower resolution images, which is ideal for our thermal image. For more details on the implementation, see Jaderberg *et al.* (2015). Once the implemented deep

learning network has been trained, it can be used to automatically classify material types as illustrated in Figure 4.5d.

The CNN architecture used the architecture in Jaderberg *et al.* (2015) as summarised in Table 4.2 and consisted of one average pooling layer, two convolutional layers, two max pooling layers and two fully connected layers. We used a Rectified Linear Unit (ReLU) as an activation function to speed up the feature learning process (Nair & Hinton, 2010). Quantised images of size 75x75 were resized to 60x60 to be feed-forwarded into the first layer to use the layer composition. Furthermore, a pilot test showed that this CNN composition did not performed good (accuracy of less than 70%). A possible reason is that the trained neural network overfitted the training set. An overfit model is more likely to produce a lower recognition accuracy on new unseen data (Tetko *et al.*, 1995). Hence, we added “dropout” to the end of the neural network layers in the CNN, which is a simple but powerful way to reduce overfitting issues (Krizhevsky *et al.*, 2012).

In the training process, a learning rate of 0.001 and a batch size of 256 were used for 350 epochs. The CNN was implemented on the *MatConvNet* framework and ran on a 64-bit Windows desktop (Intel Core i7-3770 CPU @ 3.40GHz 20GB RAM).

Table 4.2. CNN architecture used in this study which is based on Jaderberg *et al.* (2015).

Layer	0	1	2	3	4	5	6	7	8	9	10	11	12
Type	Input	Aver. Pool.	Conv.	ReLU	Max Pool.	Conv.	ReLU	Max Pool.	Fully conn.	ReLU	Fully conn.	Drop out	Soft max
Kernel size	n/a	2x2	7x7	n/a	2x2	5x5	n/a	2x2	1x1	n/a	1x1	(0.3)	n/a
Kernel Num.	n/a	n/a	12	n/a	n/a	24	n/a	n/a	48	n/a	17	n/a	n/a

Implementation of Real-Time Proximate Material Type Recogniser

Figure 4.6 illustrates the architecture of a real-time material recogniser. The smartphone client application (i) controls the mobile thermal imaging sensor (FLIR One and its SDK) to collect a temperature matrix of material, (ii) quantises the matrix, and then (iii) wirelessly sends it to a server using the TCP/IP protocol. The received data is fed to the trained deep CNN for classification. This transformer network was implemented using the *MatConvNet* framework (Vedaldi & Lenc, 2015). Alternatives based on embedded platforms with a GPU could be also

used, such as *NVIDIA Jetson TK1*¹³ with *Caffe*¹⁴. A smartphone could also be used for classification by using *TensorFlow Mobile*¹⁵

Finally, the easy-to-use API for the Quantisation and the deep learning process as well as example codes for the raw thermal data collection from the thermal camera are made available at <http://youngjuncho.com/2018/DeepThermalImagingAPI>.

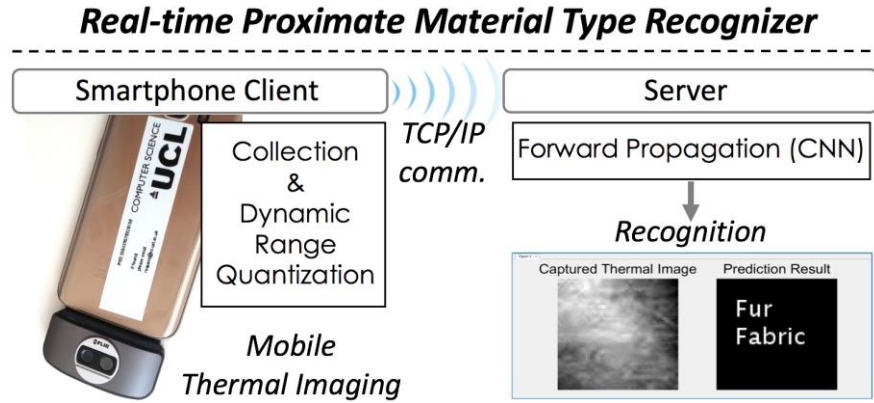


Figure 4.6. Implementation of Deep Thermal Imaging, the real-time proximate material type recogniser.

4.6.2. Data Collection Study

In this section, we report the data collection study protocol and our collected thermal image dataset. Outdoor materials we aim to collect are summarised in Figure 4.7. Outdoor materials are found in conditions which are very variable and are exposed to dirt and weathering effects. The choice of materials was linked to one of our use cases on wheelchair users discussed later. Caretakers were informally interviewed, and prior studies on wheelchair navigation were used (Kasemsuppakorn & Karimi, 2009). Materials which frequently appear on the ground (e.g., granite) or obstacles (e.g., fences) were considered important to facilitate navigation. Such materials are challenging as they are less likely to have constant patterns of spatial temperature, texture and geometry. Two materials were also added to the outdoor collection - *radiator in*

¹³ Jetson TK1: <http://www.nvidia.com/object/jetson-tk1-embedded-dev-kit.html>

¹⁴ Caffe deep learning framework: <http://caffe.berkeleyvision.org>

¹⁵ TensorFlow Mobile: <https://www.tensorflow.org/mobile>

building corridor (denoted as z) and *building entrance carpet (denoted §)* even though they are physically found indoors. The reason is that they tend to be more variable – their surfaces may be exposed to high thermal dynamics and dirt.



Figure 4.7. Outdoor materials collected for the material classification task.

To evaluate the performance of *Deep Thermal Imaging*, we conducted a study outside of controlled lab settings (e.g., temperature, light and dirt control). The recording of images was carried out by sampling while walking or running around an existing environment and taking thermal videos of material around from different perspective and repeating such recordings in various places, at different times of the day, temperature and weather conditions. This process was facilitated by the low and unsteady frame rate of the thermal camera (3-8 thermal images every second) and ensured a collection of non-similar images of the same materials to maximise the generalization capability of the trained recognition model.

The data collection process produced the dataset which includes 26,584 labelled thermal images from 17 outdoor material types. The average number of collected spatial thermal patterns from each material was 1563.8 (SD=295.3; about 300-500 images of each material per each condition). A preview video¹⁶ shows examples of data.

As illustrated in Figure 4.8, two experimenters collected thermal images in two different months (April and August) both during the day and at night. Experimenters collected spatial

¹⁶ <https://dl.acm.org/citation.cfm?id=3173576>

material temperature patterns while moving or running slowly around material during shooting to exploit the low frame rate of the thermal camera and recorded materials from different perspectives. Imaging materials from different perspectives, distances, times of days, and seasons was critical to evaluate the generalization capability of the modelling approach. However, we avoided rainy days and materials which were significantly covered by clutter such as sand or trash bags. Figure 4.9 shows examples of thermal images collected and quantised for this study. A higher number of recordings were collected for materials – such as iron fences – with a high degree of variability due to thermal reflection (Vollmer & Klaus-Peter, 2017).





		Daytime	Nighttime
<div> Data Collection Study II: Outdoor Materials </div>	In April	300-500 thermal image frames per each material 	300-500 thermal image frames per each material 
	In August	300-500 thermal image frames per each material 	300-500 thermal image frames per each material 

Figure 4.8. Data collection: thermal images of outdoor materials were recorded by two experimenters while walking/running outside both at daytime and night-time in different seasons.

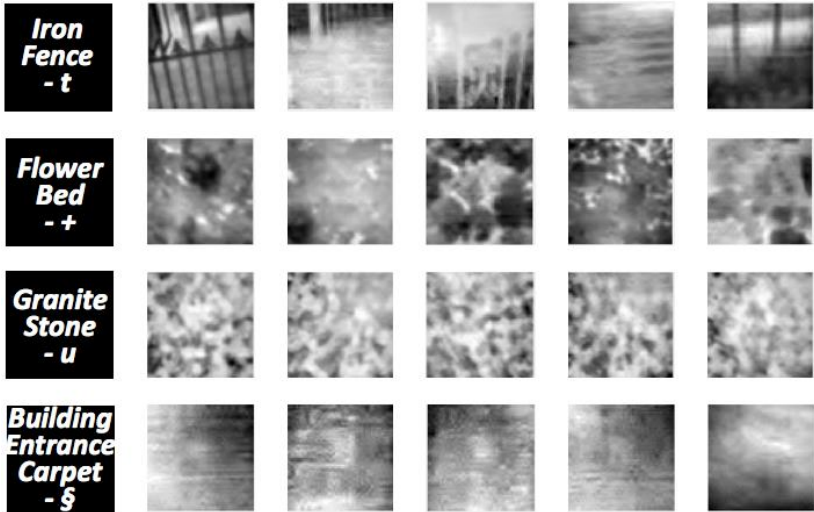


Figure 4.9. Examples of collected and quantised thermal images of t. iron fence, +. flower bed, u. granite stones (pavement edge) and §. building entrance carpet. These images contain some degree noise, e.g., dirt and rubbish.

4.6.3. Results

Figure 4.10 shows thermal dynamic levels of the collected thermal images from 17 outdoor material types. The thermal dynamics were examined by using the temperature range (T_2-T_1) in Equation (4.5). Compared with indoor materials which were additionally collected in Appendix E (Figure E.1, E.2), outdoor materials showed higher thermal dynamics (the range of thermal dynamics of indoor materials is denoted as A in Figure 4.10 and Figure E.2 for comparison). Overall, the temperatures ranged between 0.341°C and 28.744°C with a mean of 2.959°C ($\text{SD}=1.963$).

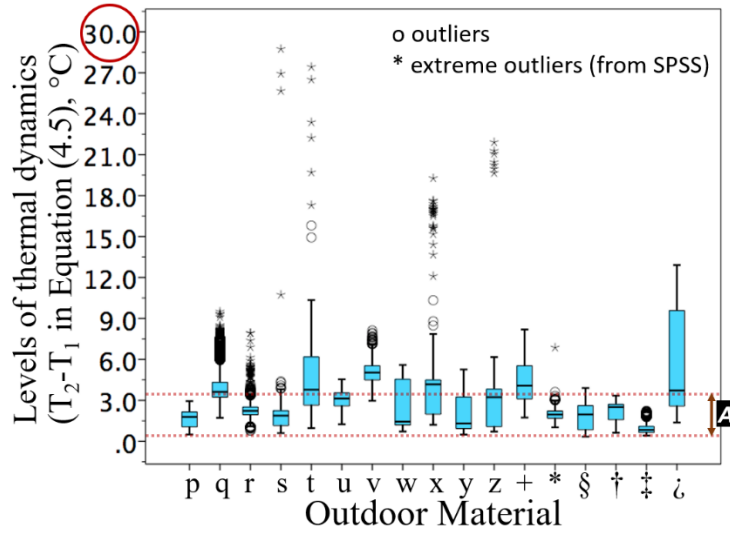


Figure 4.10. Thermal dynamics levels for the total 26584 thermal images across all 17 outdoor materials (p-⌂). The thermal dynamics were estimated by using (T_2-T_1) in Equation (4.5). The thermal dynamic range of indoor materials (in Appendix E) is marked with red dot lines (A) for comparison.

To explore the quantisation effect on material temperatures in highly variable thermal range scenes, firstly we trained the proposed classification model on the thermal dataset to which the proposed Dynamic Range Quantisation (DRQ) was not applied. For this, a fixed target thermal range of interest, $[-5^{\circ}\text{C}, 45^{\circ}\text{C}]$ was used as most temperatures of outdoor material surfaces fall into this range. Second, we trained the same classification model on the dataset which was quantised by the proposed technique.

For the training and testing, we commonly used 5-fold cross-validation. For the creation of random partitions (i.e., k-fold), the MATLAB function *cvpartition* was used. Results showed each CNN classifier produced a mean accuracy of 19.369% ($\text{SD}=21.713$) and 84.526% ($\text{SD}=9.875$) for the case **without the DRQ** support and that **supported by the DRQ**, respectively,

as shown in Figure 4.11. Figure 4.12 compares the mean class accuracy results across the targeted materials. As seen here, the *without DRQ* case showed the classification model was overfitted with the material *t* (iron fence) with some of the materials (e.g. *p*:brick walls, *z*:asphalt road with a deep pothole) not being properly classified at all.

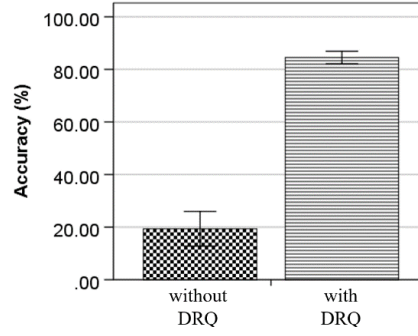


Figure 4.11. Mean class accuracy results (17 classes) from i) the proposed method, but without the Dynamic Range Quantisation (DRQ) and ii) that with the DRQ (error bar: 95% confidence interval).

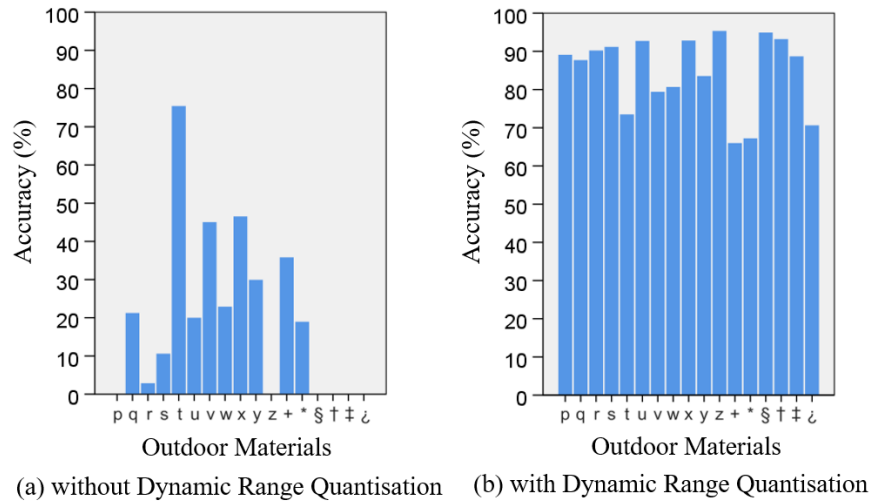


Figure 4.12. Mean class accuracy results using 5-fold cross validation: (a) without, (b) with the support of Dynamic Range Quantisation

Furthermore, we tested the DRQ supported model using a higher-fold cross-validation (10-fold) within the dataset. The aim was to understand the effects of training with smaller datasets. Using 10-fold cross validation methods, the implemented CNN classifier (Table 4.2) achieved a mean accuracy of 89.356% (SD=6.833) which was marginally higher than the result from 5-fold. Figure 4.13 summarises the mean class accuracy results from 10-fold cross-validation. Surprisingly, *z* (*radiator in building corridor*) and *§* (*building entrance carpet*), which are on the border between outdoor and indoor areas, showed highest accuracies (97.9% and

98.1%, respectively). On the other hand, + (flower bed), * (lawn) and ¿ (asphalt road with a deep pothole) still showed worst performance just below or around 80%. Those results can be explained by the fact there is a greater variability in the composition of these materials (e.g., soil, flowers).

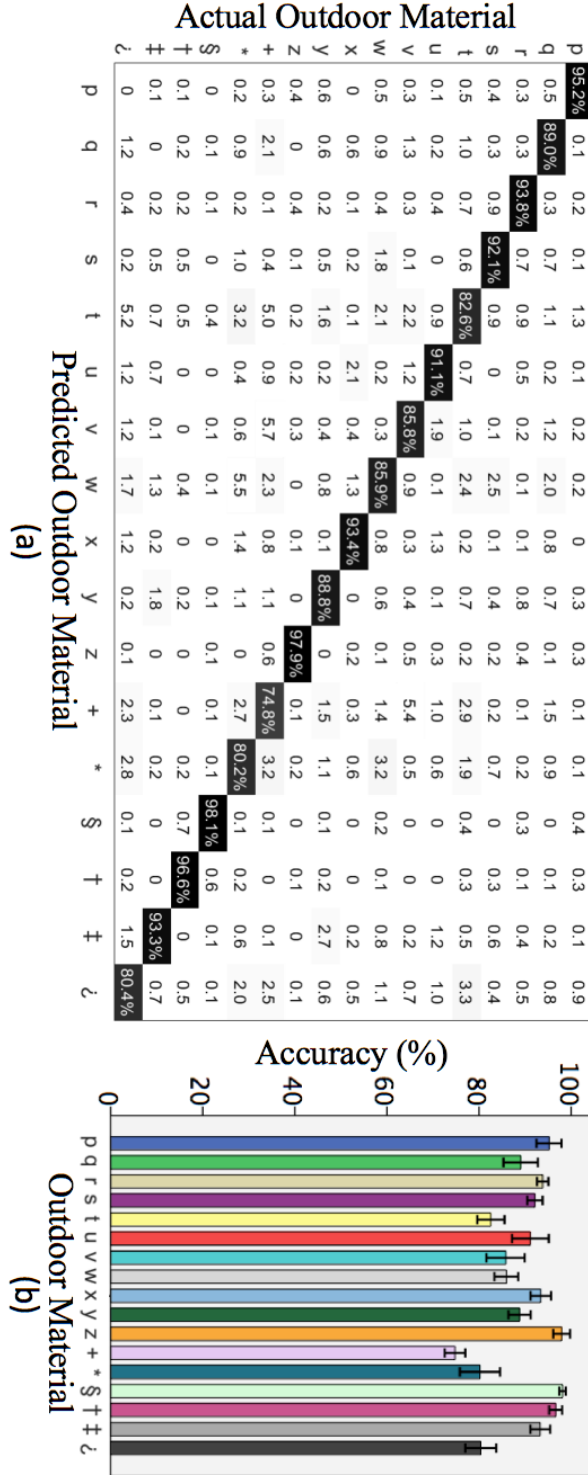


Figure 4.13. 10-fold cross validation results (*high dynamic temperature range proximate material, labels for materials shown in Figure 4.7*): (a) confusion matrix (in percentage; colour range: white(0%)-black(100%)), (b) mean class accuracy recognition distribution (error bar: 95% confidence interval).

4.7. Discussion

Given the recent greater improvements made in thermal imaging devices, helping to shrink their size, cost and weight, to the point where it can be used in much less constrained situations than previously used, we have tried to build computational methods to support the use of such sensing device in unconstrained daily, mobile situations. By focusing on the quantisation process, we have tackled the key challenge identified in Section 4.1, which addressed how thermal imaging is sensitive to environmental temperature changes. In particular, we have proposed the Optimal Quantisation technique which mainly contributes to improving the performances of existing methods in tracking skin ROIs. Through the additionally conducted study in Section 4.6, we have also demonstrated the general capability of the proposed method to different material temperatures, rather than the skin areas which are the main ROIs and an interest of this thesis.

Optimal Quantisation tackles negative effects of variable thermal range scenes on ROI tracking

Temperature of materials such as the human cutaneous skin, is affected by numerous environmental factors, in particular ambient temperatures. This leads to difficulties in selecting a temperature range for colour-mapping and feature identification which is important in either tracking ROIs or analysing spatial surface temperature patterns. However, a fixed thermal range has been generally used in the literature which explores the use of thermal imaging. To address this, we have proposed the Optimal Quantisation which adaptively adjusts the thermal range of interest along with environmental temperature effects.

An interesting advantage of the method is its ability in improving tracking performance of existing methods. To test this, we have targeted an indoor office setting. The choice of this environment is due to the mental stress-related environment in the literature, for example, the call centre in Hernandez *et al.* (2011). The indoor (but unconstrained) settings are the main environments targeted in later chapters exploring mental stress-related physiological patterns. Rather than conducting systematic analysis, we have focused on more real world-like situations, not controlling room temperatures and participants' mobilities.

As shown in Table 4.1, the state-of-the-art ROI tracking algorithms showed imperfections in tracking facial ROIs from thermal images even in this office environment with varying ambient

temperatures. On the other hand, the use of the proposed quantisation method integrated with one of the best state-of-the-art algorithms (Median Flow) significantly enhanced the performance. Given this result, we expect the capability can fundamentally improve the capability of thermal imaging in extracting physiological signals and assessing mental stress in many situations where temperature is uncontrolled.

Generalisation power of the quantisation process

As explored in Section 4.6, the quantisation process is important for a thermal investigation of not only the skin temperature but also other material temperatures in real world, outdoor environments. In particular, we have aimed to demonstrate the generalisation power of the proposed process to unconstrained, high thermal dynamic scenes far more dynamic than indoor settings. For doing so, we have collected temperatures of a variety of outdoor materials found from the wild, which have very high thermal dynamics ranging from between 0.341°C to 28.744°C as shown in Figure 4.10. Very interestingly, in the material classification tasks we have targeted here, the use of the proposed quantisation method significantly improved the classification performance of one of the state-of-the-art machine learning classifier, i.e. spatial transformer-CNN (from mean accuracy of 19.369% to 84.526%), showing the general capability of the proposed pre-processing method under different illumination, temperature conditions and point of views. The use of the *deep learning* technique also leads to a discussion of how to push its boundary by focusing on enhancing the input data in Chapter 8.

This generalisation capability has also contributed to building other HCI applications. Beyond thermal imaging, novel technological sensing approaches have been central to HCI innovations within many recent studies (Bell *et al.*, 2015; Sato *et al.*, 2015; Yeo *et al.*, 2016; Wang *et al.*, 2016). Researchers have investigated material recognition to support material awareness. Building upon them, we contribute mobile thermal imaging as a new recognition modality to supplement some of the limitations of existing approaches. Powered by the quantisation process, our approach particularly offers higher flexibility when used in the wild where there are a variety of materials and environmental factors (for example, RGB vision-based approaches are susceptible to lighting conditions).

Another benefit from this approach is that it does not require contact between the sensor and the material, reducing safety concerns, social issues and creating more seamless types of interactions and applications. We have benefited from the use of low-cost thermal cameras which

can be handheld, helping to collect thermal images from outdoors more easily. Through the author's repository website¹⁷, the collected datasets are also open to the research community.

Limitations, future opportunities and connection to next chapters

Given the main purpose of this chapter (to address environmental thermal changes), our approach does not directly handle motion artefacts which are fundamental issues in improving ROI tracking methods. Furthermore, although ambient temperatures in the indoor settings used for evaluating tracking performance in this chapter were not controlled, the level of environmental thermal dynamics in the settings was still limited. Given this, in the following chapter, we investigate how to further advance the state-of-the-art methods for automatic tracking of ROIs and respiration built on our proposed Optimal Quantisation. Furthermore, we bring thermal imaging into the real-world outdoor environments which have highly variable thermal range scenes in order to thoroughly test the capability of our advanced ROI tracking methods and make mobile thermal imaging more feasible in HCI applications.

4.8. Chapter Summary

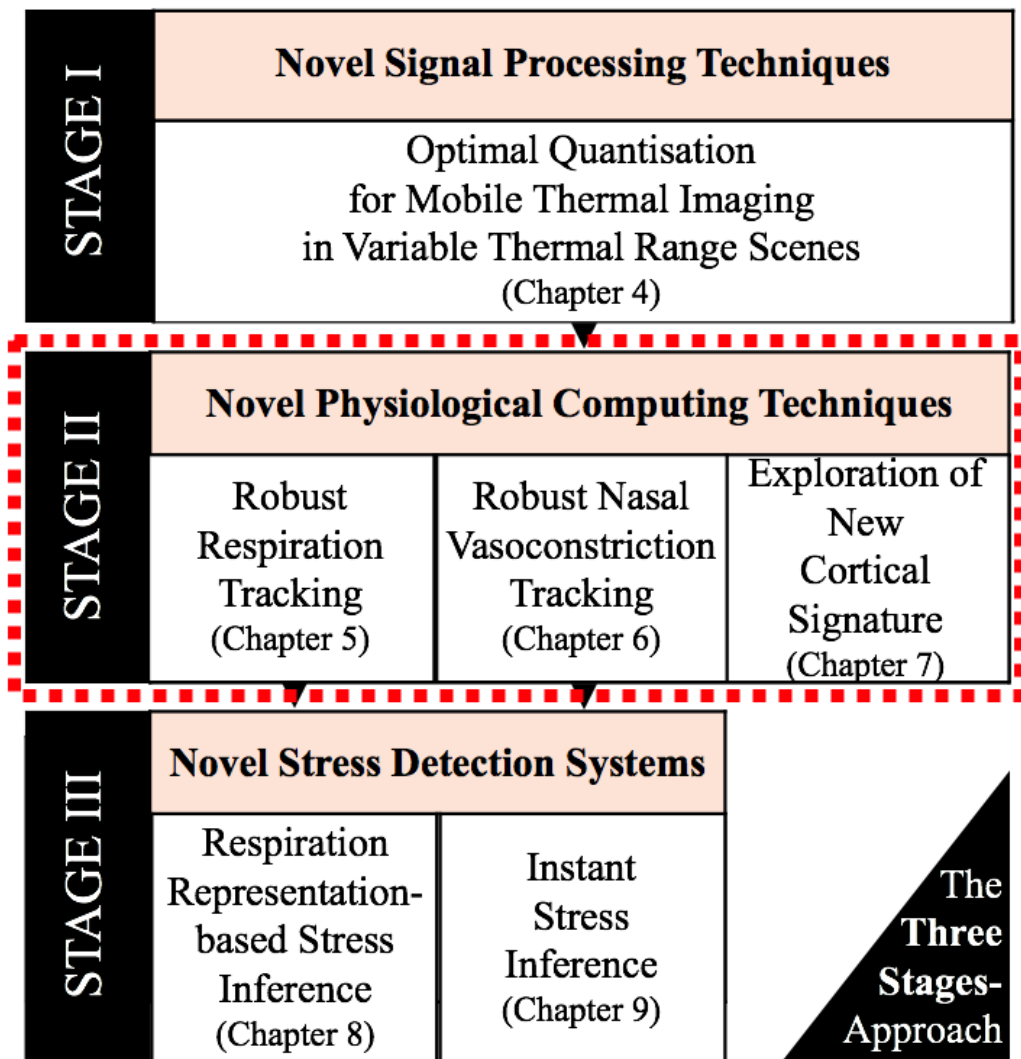
In this chapter, we have proposed a novel thermal image processing technique, *Optimal Quantisation*, which continuously adapt colour mapping of thermal data to ambient temperature dynamics, in turn compensating for negative effects of variable thermal range scenes. The main contribution of the proposed Optimal Quantisation is two-fold. Firstly, this method has significantly improved performance of the state-of-the-art algorithms in tracking the skin ROIs on thermal images collected from indoor unconstrained settings without constraining room temperatures and participants' mobility. Secondly, our proposed technique has enabled the latest deep learning algorithm to properly analyse and classify other material temperatures (than human skin temperatures). To the best of our knowledge, this work provides the first demonstration that spatial thermal patterns of materials can be used to enable automatic material recognition tasks with very good performance (above 84% mean accuracy from 17 outdoor materials, which is significantly higher than the model without our quantisation method:19.369%).

¹⁷ <http://youngjuncho.com/datasets>

We have leveraged a mobile, low-frame rate thermal camera to collect datasets in unconstrained settings. In particular, this helped to create a large dataset (26,584 images in Section 4.6, and additionally, 14,860 images in Appendix E) of variant thermal information of materials. All the collected datasets and API built in this chapter have been released to foster this research.

STAGE II

Novel Physiological Computing Techniques



Chapter 5

Robust Respiration Tracking in Unconstrained Settings

In this first part of Stage II, we aim to build a computational method for reliably estimating respiration patterns through mobile thermography in unconstrained settings. Major challenges concerned here are how to improve the performance of the nostril ROI-tracking in real world-like unconstrained situations and how to reliably extract respiration cues from the tracked ROI sequence given the low signal quality of mobile thermal imaging. Accordingly, two computational methods are proposed in this chapter: Thermal Gradient Flow for nostril-ROI tracking and Thermal Voxel Integration for estimating respiration patterns. The proposed techniques are thoroughly evaluated using datasets collected from both controlled and unconstrained situations.

Section 5.1 discusses background and our research question. Section 5.2 proposes the main computational methods. Sections 5.3 and 5.4 describe experimental protocols, datasets and evaluation methods, which are followed by Section 5.5 where evaluation results are presented. The chapter ends with a discussion and a summary in Section 5.6 and 5.7, respectively.

5.1. Background and Research Question

Monitoring respiratory rate plays a key role in a range of applications that span from direct diagnosis of and treatment for lung problems (e.g. hyperventilation, apnoea and interstitial lung disease) and cardiovascular conditions to supporting a person's psychological needs (e.g. stress, anxiety regulation) (Grossman, 1983; Everly Jr & Lating, 2012; West, 2012). Despite its importance, it has been largely disregarded in real world healthcare technology applications

(Cretikos *et al.*, 2008). One possible reason is the inconvenience of conventional respiration measurement systems, such as chest-belts or oronasal probes, which demand direct physical contact as discussed in the literature (Kumar *et al.*, 2015; Pereira *et al.*, 2015; Gastel *et al.*, 2016). These systems are often uncomfortable to wear and prone to motion artefacts which might cause incorrect sensor readings. In addition, in some medical and chronic conditions, where monitoring everyday physiological processes may be pivotal, direct contact with the skin may not be acceptable, such as Complex Regional Pain Syndrome (Birklein, 2005).

Contactless ways to measure respiration-related signatures can help address these limitations. They are remote-photoplethysmography (PPG) (Poh *et al.*, 2011; Gastel *et al.*, 2016), Doppler radar (Droitcour *et al.*, 2009) and thermal imaging (Murthy & Pavlidis, 2006; Fei & Pavlidis, 2010; Abbas *et al.*, 2011; Lewis *et al.*, 2011; Pereira *et al.*, 2015) based measurements. However, they have not been investigated in the context of ubiquitous and mobile computing situations.

With the assistance of ambient lights (e.g. natural sunlight, lamp), a digital image sensor such as an RGB-based camera can be used as a remote PPG sensor (Verkrusse *et al.*, 2008; Poh *et al.*, 2011; Xu *et al.*, 2014; Kumar *et al.*, 2015; Gastel *et al.*, 2016) for monitoring blood volume pulse related parameters (e.g. heart rate). Poh *et al.* (2011) and Gastel *et al.* (2016) found that PPG could also detect periodic respiratory periodic cycles in a known respiratory rate range (10-40 BPM in Gastel *et al.*, 2016), opening the potential of real-time contactless breath tracking in stationary environments. However, although the sensors for PPG can be mobile, fundamental issues with PPG itself limit its flexibility and mobility. In particular, RGB-based cameras rely on moderate, stable ambient light levels. Therefore, they cannot easily support physiology monitoring in extreme lighting conditions which are either very dark or very bright. Furthermore, they struggle in situations where the light conditions continually change. As a result, most research with remote PPG has been applied in controlled conditions where the ambient lighting levels can be controlled (e.g. a constant 500 lux brightness in Kumar *et al.*, 2015).

Other approaches can be used which are immune to ambient lighting conditions. Active sensors such as radar provide their own illumination. However, for respiratory rate tracking (such as Droitcour *et al.*, 2009) they are restricted to monitor physiological parameters in stationary settings and indoors, where it is easier to ensure stillness of both the person and of the hardware installation, limiting their application in real-world deployments.

On the other hand, thermal imaging does not have many of these constraints. As reviewed in Chapter 2, thermal imaging interprets the electromagnetic radiation which is naturally emitted

from any object; hence, it can measure temperature in a passive way and does not require lighting sources. Work on thermography has shown it is possible to track respiration in a contactless manner by monitoring the temperature changes around the nostrils which are caused by inhalation and exhalation breathing cycles (Murthy & Pavlidis, 2006; Fei & Pavlidis, 2010; Abbas *et al.*, 2011; Lewis *et al.*, 2011; Pereira *et al.*, 2015). However, even state-of-the-art methods in tracking respiration through thermal imaging require to constrain a person’s mobility (Lewis *et al.*, 2011; Pereira *et al.*, 2015). More importantly, all the methods depend on the use of large-sized, immobile thermal imaging systems demanding large indoor spaces for the hardware installation, similar to radar systems. This leads to the research question: *Can respiratory signatures be recovered through thermography in unconstrained HCI settings? If so, how can we build methods for robust respiration tracking using thermal imaging?*

With the use of low-cost, lightweight and very small thermal cameras, we aim to push the boundary of thermal imaging-based respiration tracking by bringing them into mobile situations. However, as discussed in Chapter 4, mobile thermal imaging is also challenged in scenes with high thermal dynamic ranges (e.g. due to the different environmental temperature distributions indoors and outdoors). This challenge is further amplified by general problems such as motion artefacts and low spatial resolution, leading to unreliable breathing signals. Below are detailed descriptions of such challenges.

Challenge I: Difficulty in Tracking the Nostril under Motion Artefacts

A recent body of work on thermal imaging and remote-PPG has employed motion-tracking algorithms to extract physiological features from a ROI under motion. For example, Kumar *et al.* (2015) and Gastel *et al.* (2016) used the Lucas-Kanade (KLT) algorithm (Lucas & Kanade, 1981; Shi & Tomasi, 1994) to track facial areas where the PPG signal can be extracted. Pereira *et al.* (2015) applied Mei *et al.* (Mei & Ling, 2011)’s Sparse Representation-based tracker. Although the body of work has achieved high performance of the tracking of ROIs in indoor constrained situations where just a small amount of a person’s head motion is allowed, unconstrained motion artefacts have not been generally tested.

In particular, the thermally expressed nostril shape can undergo a significant amount of deformations. These deformations result mainly from a person’s movements. For example, the deformations shown in Figure 5.1a result from the participant turning their head and laughing. Therefore, although static contexts have been the main target for research, it is not possible to

guarantee, in real-world settings, that a person's nostril will be in a fixed, known shape. Therefore, new tracking algorithms must be developed which are robust to deformation of ROIs.

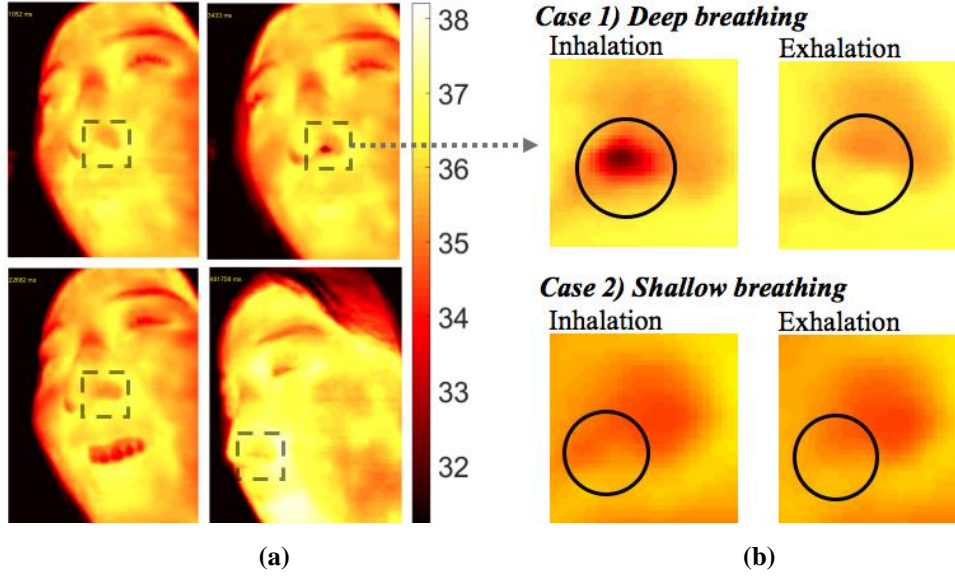


Figure 5.1. Key challenges in respiration tracking through thermal imaging in real-world settings: (a) difficulty in tracking the nostril ROI: the shape of the nostril is affected by one's mobility and thermal dynamics, (b) low signal quality of respiration patterns: four examples of thermal images show the tracked nostril region while breathing. The traditional average temperature is not ideal in extracting respiratory patterns when the respiration-driven thermal variance is weak, e.g. during shallow breathing (Case 2) compared with deep breathing (Case 1). The low spatial resolution of mobile thermal imaging also leads to the weak signal.

Challenge II: Low Quality of Respiratory Signals from Mobile Thermal Imaging

The most common way to track respiration rate is to analyse sequences of an *average* temperature in a tracked nostril region, which fluctuate caused by the expiration and inspiration cycles (Murthy & Pavlidis, 2006; Fei & Pavlidis, 2010; Abbas *et al.*, 2011; Lewis *et al.*, 2011; Pereira *et al.*, 2015). However, in many cases, the temperature change associated with breathing can be fairly small and affect only a small number of pixels. Figure 5.1b for example, illustrates the difference between shallow and deep breathing. Another difficulty is that the average temperature can be strongly affected by subtle location changes of the ROI bounding box and windy situations which can cause sudden global changes in the temperature distribution. A final difficulty is that when a person's viewing direction changes, it can also decrease the number of pixels that contain respiratory information. This is even more important for mobile thermal imaging systems, which often have low spatial resolution. Consequently, for the enhancement of

the respiratory signal quality, there is a need to design a new computational method which is robust to the various factors as discussed above.

5.2. Proposed Computational Methods

To overcome the challenges mentioned above, a key technical strategy is proposed here. The strategy is built on the Optimal Quantisation proposed in Chapter 4 (i.e. adaptive conversion of two-dimensional temperature matrixes into colour-mapped thermal images reflecting different thermal dynamic range scenes). Figure 5.2 illustrates the overall process. As the process for the extraction of respiratory signs from thermal imaging sequences, the proposed method consists of (i) Thermal Gradient Flow – a nostril tracking algorithm to reduce effects of motion dynamics and (ii) Thermal Voxel-based Respiration Estimation to enhance the respiratory signal quality. Details of each component are described below.

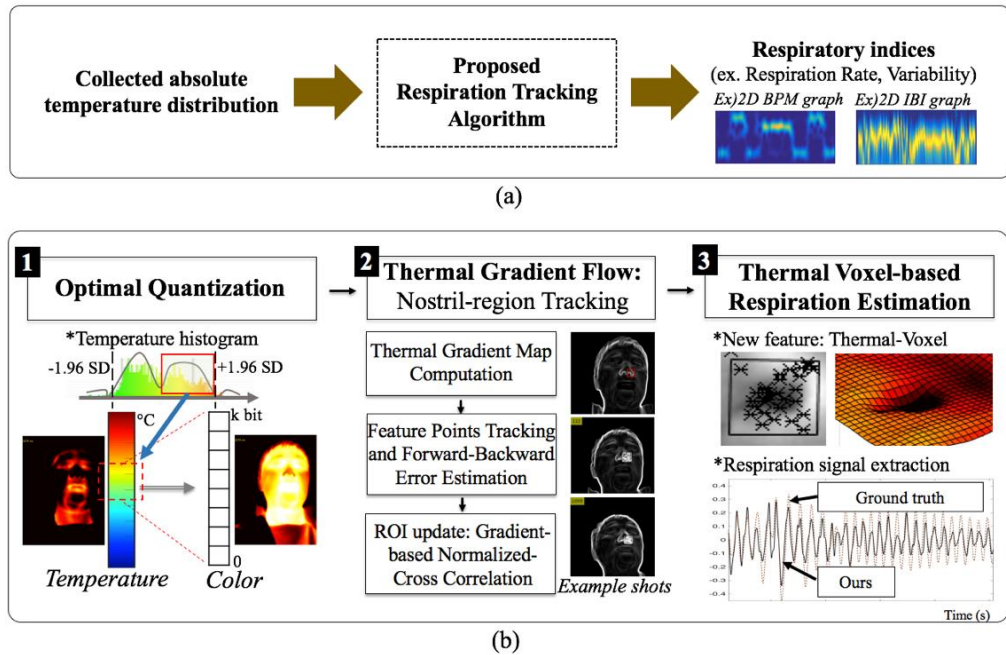


Figure 5.2. (a) Overall procedure of thermal imaging-based respiration tracking, (b) key components: 1) Optimal Quantisation (proposed in Chapter 4), 2) Thermal Gradient Flow – nostril-region tracking method based on the thermal gradient magnitude computation and points tracking methods, 3) Thermal Voxel-based Respiration Estimation – extracting the respiratory signals by integrating the unit thermal-voxels inside the nostril.

5.2.1. Robust Nostril-ROI Tracking: *Thermal Gradient Flow*

To achieve robust performance in the tracking of the nostril, we propose a new algorithm called *Thermal Gradient Flow*. Given the colour-mapped images produced through the Optimal Quantisation, this algorithm computes the thermal gradient magnitude matrix for each frame and employs the Median Flow algorithm (Kalal *et al.*, 2010) which uses forward-backward error estimation on points tracked by Lucas-Kanade's disparity-based tracker (Lucas & Kanade, 1981). To enhance robustness, we compensate for the loss of feature points by resetting a ROI based on the gradient-based two-dimensional normalised cross correlation, as illustrated in Figure 5.2b-2.

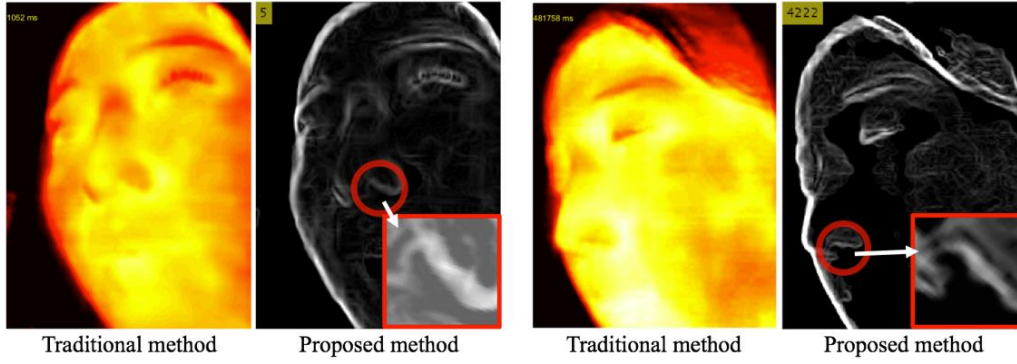


Figure 5.3. Example shots of conversion from thermal images to thermal-gradient magnitude maps: the proposed method can help to preserve the morphology of the nostril region during motion (Zoomed-in-areas are rotated for the visual representation).

The human homeothermic metabolism and the relatively low thermal conductivity of the human skin act as low pass filters. As a result, the shape of nasal and nostril areas is often blurred, leading to a weak differentiation between key facial features which are required to support ROI trackers (Shi & Tomasi, 1994). To obtain clearer features, we enhance the boundary between the nostril and the ala of the nose by converting the quantised thermal image u into the two-dimensional thermal-gradient magnitude map Φ :

$$\Phi(x, y) = \sqrt{\left(\frac{\partial u(x, y)}{\partial x}\right)^2 + \left(\frac{\partial u(x, y)}{\partial y}\right)^2} \quad (5.1)$$

where x and y are the coordinates in the x - y plane of the image space. As illustrated in Figure 5.3, the more distinct morphological shape of the nostril can be obtained from the thermal-gradient

based image than the normal thermal image under combined artefacts (i.e. motion and respiration dynamics). The converted image can then be used to collect the positions of features representing the nostril for the motion-tracking.

The thermal-gradient map of the nostril-ROI, Φ_{ROI} , is chosen by selecting the nostril as a ROI of size of $N \times N$ pixels in the first frame. In this thesis, the initial selection of ROIs is done manually. As our point tracker, we use the *Median Flow* algorithm (Kalal *et al.*, 2010) because it has proven its tracking performance on thermal imaging in a non-biomedical sector (Abouelenien *et al.*, 2017). This algorithm calculates the forward-backward error defined as

$$e(T_f^k | S) = \|x_t - \hat{x}_t\| \quad (5.2)$$

where the thermal-gradient image sequence is $S = (\Phi_t, \Phi_{t+1}, \dots, \Phi_{t+k})$, the forward trajectory is $T_f^k = (x_t, x_{t+1}, \dots, x_{t+k})$ and the backward trajectory $T_b^k = (\hat{x}_{t+k}, \hat{x}_{t+k-1}, \dots, \hat{x}_t)$ produced by backward tracking up to the initial frame t (and $t+k$ is the current frame). Here, $\hat{x}_{t+k} = x_{t+k}$ and $\|x_t - \hat{x}_t\|$ is the Euclidean distance between the two points (facial features). In this algorithm, the points are tracked using the KLT (Lucas & Kanade, 1981). In our case, the points are selected from the nostril ROI on the thermal-gradient magnitude map. For more details on the implementation, we refer to the *Median Flow* (Kalal *et al.*, 2010).

As a final strategy to handle the case when the features from facial areas are completely lost, we can use two-dimensional normalised-cross correlation (Lewis, 1995). In particular, the combination of correlation with gradient has been shown to give high performance in the registration of deformable components in neuroimaging (Avants *et al.*, 2008, 2011). Similarly, we enhance the tracking performance of the nostril by searching for a new position of the ROI which maximises the *gradient-based normalised-cross correlation coefficient*

$$\gamma(\mathbf{x}) = \frac{\sum_{\tilde{\mathbf{x}}} (\Phi_{ROI}(\tilde{\mathbf{x}}) - \mu_{\Phi_{ROI}(\mathbf{x})}) (\Phi(\tilde{\mathbf{x}} - \mathbf{x}) - \mu_{\Phi})}{\sqrt{\sum_{\tilde{\mathbf{x}}} (\Phi_{ROI}(\tilde{\mathbf{x}}) - \mu_{\Phi_{ROI}(\mathbf{x})})^2 \sum_{\tilde{\mathbf{x}}} (\Phi(\tilde{\mathbf{x}} - \mathbf{x}) - \mu_{\Phi})^2}} \quad (5.3)$$

where \mathbf{x} is a set of (x, y) in Equation (5.1) which is situated at the centre of $N \times N$ square (i.e. same with the size of a ROI), $\tilde{\mathbf{x}}$ is another set of (x, y) which falls in the region under the ROI at

the previous frame, $\mu_{\Phi_{ROI}}(\tilde{x})$ is the mean of $\Phi_{ROI}(\tilde{x})$. When the number of tracked points falls below a certain threshold, this method resets the ROI and finds new gradient-based point features. This method can also be applied to the automated ROI selection at the initial frame in case we have a plenty of nostril-image sets.

5.2.2. Respiration Estimation by *Thermal Voxel* Integration

Heat exchange in nostrils during inhalation and exhalation is determined by a person's breathing pattern. The most frequently used method to represent breathing signals is the *averaging* of temperatures inside the nostril ROI (Murthy & Pavlidis, 2006; Fei & Pavlidis, 2010; Abbas *et al.*, 2011; Lewis *et al.*, 2011; Pereira *et al.*, 2015). By applying this to every single thermal image frame, we can extract a breathing-related time series.

However, there are critical issues with the use of the averaging method. Therefore, in this section, we present a new three-dimensional *Thermal Voxel*-based feature to further enhance the quality of the breathing signals. Inspired by the use of voxels in neuroimaging (Ashburner & Friston, 2000), the proposed approach maps each two-dimensional thermal element into a thermal voxel in a virtual three-dimensional space as visualised in Figure 5.4a and 5.4b. This feature can be immune to issues regarding low quality of breathing information and global changes of spatial thermal distribution since it can help to extract breathing-induced thermal volume changes inside the nostril by computing the quantity of inside thermal voxels as shown in Figure 5.4b. It can be constructed from

$$\hat{v}(t) = \int_{T_{\min}(t)}^{T_{\delta}(t)} \Lambda_t(T) dT \approx \sum_i \sum_j T_{\delta}(t) - \hat{u}_{ij}(t), \quad T_{\delta}(t) - \hat{u}_{ij}(t) > 0 \quad (5.4)$$

where $\Lambda_t(T)$ is the integral of the thermal voxels in the nostril in a cross section with the temperature T , \hat{u}_{ij} is the absolute temperature on the tracked region and T_{δ} is the upper boundary to integrate the concave volume which is set to a temporal moving average (here, $n=2$ to which is empirically set) of the spatial mean temperature values to not only have a stable boundary but to consider global thermal changes. If the ROI tracker loses track of the nostril region, T_{δ} would need to be reset from the next frame to reject its value from the misplaced

bounding box. The misplacement could be detected by inspecting a sudden change (i.e. differential) of statistical skewness on the thermal distribution:

$$\Delta\gamma_1^t = E\left[\frac{\hat{u}_{ij}(t) - \mu_t}{\sigma_t}\right]^3 - E\left[\frac{\hat{u}_{ij}(t-1) - \mu_{t-1}}{\sigma_{t-1}}\right]^3. \quad (5.5)$$

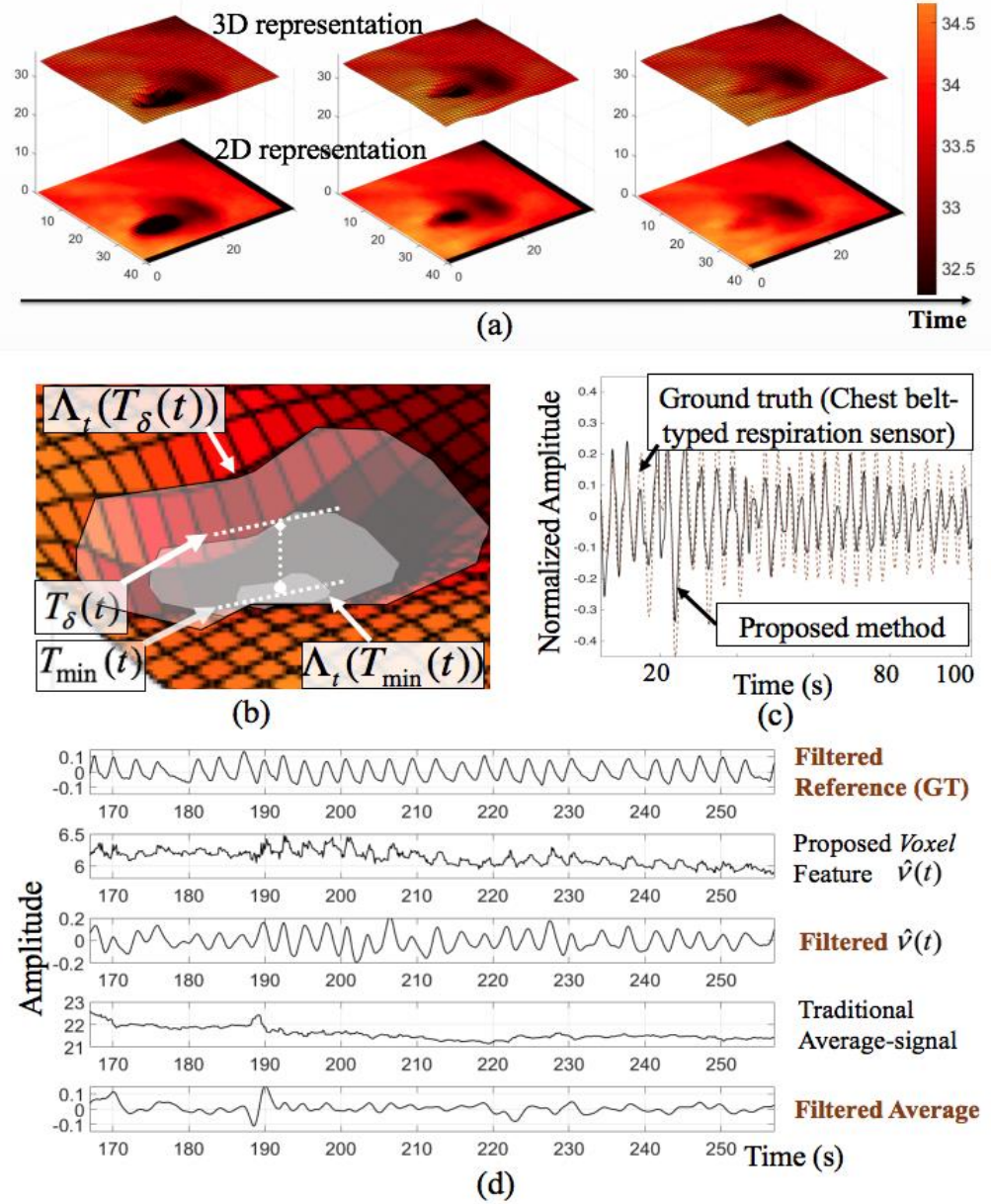


Figure 5.4. Extraction of respiratory patterns through Thermal Voxel integration: (a) a person's nostril and its thermogram sequences along the time in 3D (top) and 2D (bottom), (b) the concave volume corresponding to heat variances in the nostril, (c) the extracted respiratory signals compared with ground truth signals, and (d) a comparison of the filtered voxel-based signals with the traditional method as the participant changes their head. The voxel method closely tracks ground truth, but the traditional method fails.

To estimate respiratory rate, both frequency domain approaches, such as peak frequency detection (Gastel *et al.*, 2016), and time-domain approaches, such as the Bayesian approach based on short-time estimators (Brüser *et al.*, 2013) employed in (Pereira *et al.*, 2015), could be used. In this work, we used the short-time power spectral density, which analyses the self-similarity of the thermal voxel feature $\hat{v}(t)$ to determine the rate. This is achieved by computing the Fourier transform F_f of the short-time autocorrelation function (Gray & Davisson, 2004). To decrease ripples in the frequency domain due to the truncated short-time window, we use the Gaussian window

$$w_i(k) = \hat{v}(n_i + k) e^{-\frac{1}{2} \left(\frac{k}{\sigma} \right)^2}, \quad k \in \{-\hat{t}_{\max} f_s, \dots, \hat{t}_{\max} f_s\}. \quad (5.6)$$

The window length $w_i(k)$ is $2\hat{t}_{\max} f_s + 1$, where \hat{t}_{\max} is the upper time limit of respiration and f_s is the sampling frequency. The value of \hat{t}_{\max} is determined by the expected minimum respiratory rate of interest. For our experiments, we use the same range of expected breathing rate as in (Pereira *et al.*, 2015) of 0.1Hz to 0.85Hz. Therefore, $\hat{t}_{\max} = 10$ s. Once $w_i(k)$ has been computed, its value is normalised by feature scaling. The output is filtered through a third order elliptic filter (a passband ripple of 3 dB and a stopband attenuation of 6 dB) with passband cut-off frequency of 0.1Hz and 0.85Hz. Given this, we estimate the respiratory rate by searching for frequency f which maximises the power spectral density

$$S_V(f) = F_f(R_{ww}) = \sum_k R_{ww}(k) e^{-j2\pi f k} \quad (5.7)$$

where R_{ww} is the short-time autocorrelation of the filtered $w_i(k)$.

The proposed technical components are thoroughly evaluated in both laboratory and real-world settings using the datasets produced from the experimental protocol in Section 5.3. Accordingly, the contribution made in this chapter is a robust respiration tracking method which can support both the controlled and unconstrained settings.

5.3. Experimental Protocols and Datasets (*ThermalBreath I, II*)

To evaluate the proposed methods, we aim to gather thermal images which contain respiration information from situations with different levels of complexity in terms of thermal dynamics and motion artefacts. With this purpose, we target three situations:

Situation 1) controlled respiration in environments with non-constant temperature

Situation 2) unconstrained respiration during desk activity with natural motion artefacts

Situation 3) unconstrained respiration in fully mobile context and varying thermal dynamic range scenes

As we can notice here, *Situation 2* is where we have collected the *ROITracking* dataset in Chapter 4. In addition to this, we have conducted two different experiments to create two datasets along with the first and third situations. In total, three datasets are used in the evaluation process.

ThermalBreath I: Controlled respiration in environments with non-constant temperature (Situation 1)

The aim of this experiment is to carry out a systematic evaluation of our approach. 5 healthy adults (2 female) (aged 29-38 years, $M=31.4$, $SD=3.78$) were recruited from the university subject pool. Following the protocol used in studies on non-contact breathing monitoring (Pereira *et al.*, 2015; Gastel *et al.*, 2016), participants were asked to maintain a stable posture and breathing according to a set of predefined breathing patterns presented to them on a screen in an indoor environment with controlled room temperature. Unlike the protocols in the literature, we repeated the same task with the breathing exercise in different environments with dynamically varying ambient temperature. This is to systematically explore the effect of variable environmental temperatures on ROI tracking as well as respiration tracking.

Figure 5.5 shows the design for this experiment. As shown in Figure 5.5a, all the participants were given a thermal camera attached to an Android smartphone to record their facial temperatures and an additional smartphone that provided the breathing patterns (guide). Figure 5.5b shows the three guiding breathing patterns composed of slow (10 breaths/min), normal (15 bpm) and fast speed (30 bpm). Each breathing pattern lasted for 30 seconds. The guiding breathing patterns were displayed dynamically on the screen. Participants were given a 60 seconds-training period. Taking advantage of mobile thermal imaging, participants were able to

monitor themselves by aiming the camera at their face. The distance between the face and the device ranged from 35cm to 55cm. Four different places in which the task was repeated are: a controlled room (“Place A”), entrance of the building (wind from outside and heat from inside) (“Place B”), a street corner (windy) (“Place C”) and park (“Place D”). Each person was asked to remain as still as possible. All sessions were run in winter. The collected dataset (*ThermalBreath I*) consists of approximately 80 minutes recordings (5 participants x 4 places x 4 minutes).

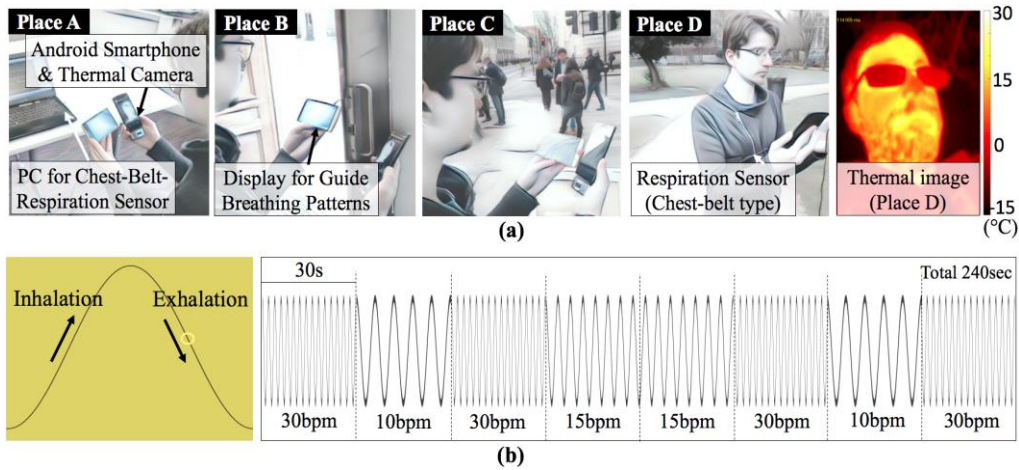


Figure 5.5. Experiment with controlled respiration in environments with non-constant temperature: (a) to obtain different thermal dynamic range scenes (i.e. environments with non-constant dynamic temperature), four different places were chosen (Place A: room, B: entrance of the building, C: corner on the street, D: park), the last image-shot is a thermal image collected in Place D, the experiment was run in winter, (b) the guiding breathing patterns are composed of three different rates (10(slow), 15(normal), 30(fast) breaths/min).

ThermalBreath II: unconstrained respiration in fully mobile context and varying thermal dynamic range scenes (Situation 3)

The last experiment aims to measure the respiration patterns from people undertaking natural, unrestricted actions. We recruited 8 healthy adults (5 female), aged 23-31 years ($M=27.0$, $SD=2.93$) from various ethnical backgrounds. In order to enable mobility, the thermal camera was attached to a headset-microphone-shaped mount whose distance from the face ranged between 20cm and 30cm as shown in Figure 5.6 (an example of collected thermal images is also shown here). To simulate a variety of fully unconstrained situations, the experiment had two main sessions: i) indoor physical activity and ii) outdoor physical activity.

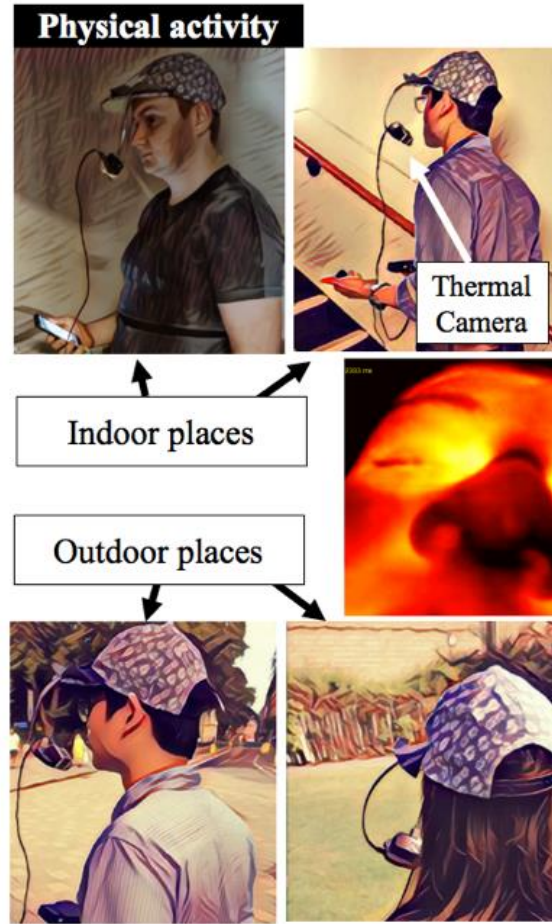


Figure 5.6. Experiment with unconstrained respiration in both indoor and outdoor light physical activities.

The first session consisted of three tasks of 2 minutes each: walking through a corridor, standing in a dark room while having small movements and climbing and descending stairs. The second session was carried out outdoor on a street pavement and in a windy park to involve varying thermal dynamic range scenes. During the session, participants were guided to walk slow, walk fast, and stroll in natural paces. Each walking pattern lasted for 2 minutes. All sessions were run in summer. The final dataset (*ThermalBreath II*) includes thermal imaging sequences of approximately 96 minutes (8 participants x 2 sessions x 3 activities x 2 minutes).

Note that, following Gastel *et al.* (2016), only 5 participants were invited for collecting the first dataset (*ThermalBreath I*). Given the very controlled nature of this experiment, this number was considered sufficient to ensure the robustness of the test. On the other hand, a slightly higher number was used since higher variability between participants breathing patterns was expected for the other two situations (the dataset *ThermalBreath II* here and *ROITracking* in Chapter4). Due to ethics restrictions, only mild physical activity was used at this phase and only

healthy participants were invited. The slightly higher number was considered sufficient to cover the expected amount of variability.

In collecting all three independent datasets (*ThermalBreath I, II* from here, *ROITracking* from Chapter 4), a commercial respiration sensor was used to evaluate the performance of our approach for respiration tracking. Participants were required to wear a chest belt-based respiration sensor (ProComp Infiniti Resp/SA9311M, Thought Technology). This reference respiration sensor produces the respiration waveform by monitoring expansion and contraction of the chest or the abdomen. The reference sensor collects the data at 256Hz. Therefore, to allow a direct comparison, we up-sampled the data sequences from the Thermal Voxel-based method with spline interpolation to 256Hz. Additionally, we also collected blood volume pulse using a finger PPG sensor (ProComp Infiniti BVP/SA9308M, Thought Technology) with the purpose to build a rich physiological dataset

The ROI sequences which were completely tracked by Optimal Quantisation and Thermal Gradient Flow from all the datasets are publicly available.

5.4. Evaluation Method

Evaluating ROI-tracking performance

Three state of the art visual tracking algorithms proposed between the 1990s and 2010s were implemented to evaluate the ROI-tracking performance of the Optimal Quantisation based Thermal Gradient Flow method. As in the previous chapter which explores how the Optimal Quantisation can boost the performance of existing tracking methods, we implemented the Mode-seeking (also called mean-shift) (Cheng, 1995) and Median Flow (Kalal et al., 2010).

With the aim to compare the proposed method with the state-of-the-art method in thermal imaging-based respiration tracking (Pereira et al., 2015), we further implemented Mei et al.'s Sparse Representation (Mei & Ling, 2011) which is the backbone of Pereira et al.'s latest work. As in Chapter 4, the benchmark algorithms were built on the dominantly used quantisation technique that takes into consideration temperatures only within a fixed thermal range of interest, e.g. [28°C, 38°C] in Fei and Pavlidis (2010).

In most cases, the algorithms stopped working when a tracking fault occurred, i.e. tracker going off beyond approximately 50% of the nostril region. We also manually checked the tracking faults to confirm the number of fully tracked frames. This number was used to describe tracking performance levels. For the parameter settings with both the Median Flow and the Thermal Gradient Flow, the maximum backward-forward error allowed (i.e. Equation (5.2)) was set to 5 for *ThermalBreath I* and *ROITracking* datasets (i.e. ‘small’ ROI), and to 25 for the *ThermalBreath II* dataset (i.e. ‘big’ ROI). The setting of the values was also based on the physical distance between the camera and the nostril.

As in the previous chapter (Section 4.5), we used Friedman’s analysis of variance (ANOVA) test to investigate the statistical difference of means of ROI-tracking results for each algorithm type. By contrast with Situations 2 and 3 (*ROITracking* and *ThermalBreath II*) where the type of environments did not vary substantially across each case, the first situation (*ThermalBreath I*) included multiple environmental temperature settings; hence, we carried out the significance test separately for the indoor (i.e. indoor room Place A) and outdoor groups (i.e. outdoor places: Place B-D). Lastly, the Friedman one-way ANOVA was followed by the Wilcoxon signed ranks test for post hoc pairwise comparisons.

Evaluating respiration rate tracking performance

To evaluate tracking of respiratory rate, the Thermal Voxel based respiration estimation method was compared with the standard and state-of-the-art approach of using the average temperature over the nostril ROI (Murthy & Pavlidis, 2006; Fei & Pavlidis, 2010; Abbas *et al.*, 2011; Lewis *et al.*, 2011; Pereira *et al.*, 2015). In this evaluation, we used the ROI-sequences that were automatically tracked by Optimal Quantisation-based Thermal Gradient Flow tracking method. The proposed and benchmark methods were implemented in MATLAB (2015b, The MathWorks). The validation process was carried out on a 64-bit Windows 7 desktop (Core i3-4160T 3.10GHz process, Intel) with 8GB RAM.

As reference signals, the waveforms collected from the chest-belt-respiration sensor were used as described in Section 5.3. To enable the comparison of the two sensed signals, these were automatically synchronized using the *Maximum-Amplitude of Cross-Correlation (MACC) alignment* expressed as $\max(R_{fg}(\eta))$ where R_{fg} is the cross-correlation between the reference signals and the estimated signals, and η is the discrete lag. As shown in Figure 5.7a, we synchronized the two signals by analysing periodic similarity. Even though periodicity changes

throughout the experiment, correct alignment could be achieved when the cross correlation is applied across all the data for a single trial. Here, the length of the time-window (i.e. Equation (5.6)) was set to 20 seconds with 15 seconds overlap (i.e. 75%).

To statistically compare two different respiratory rate measurements, we used the Bland-Altman plot and the root mean square error (RMSE) as in Kumar *et al.* (2015). In particular, the Bland-Altman method is to analyse the agreement between two different methods, compensating for the correlation analysis (Altman & Bland, 1983).

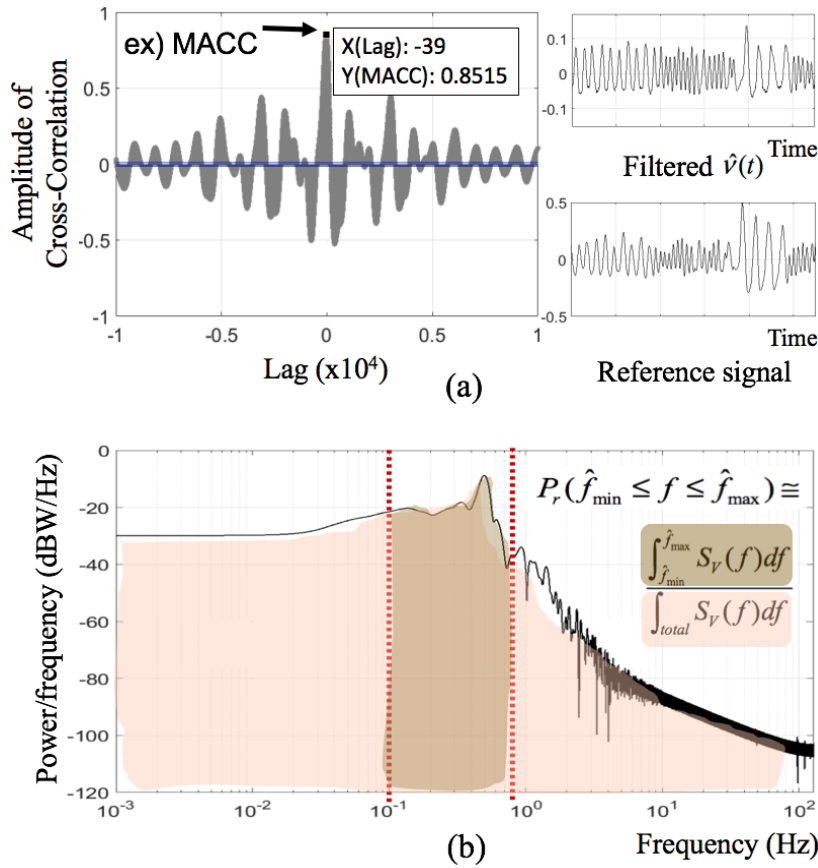


Figure 5.7. Statistical methods for evaluations: (a) automated-synchronization between estimated signals and reference signals using the *maximum-amplitude of cross-correlation (MACC)*, (b) respiration-related goodness probability as a respiratory signal quality index (SQI).

To examine the respiration signal quality, we use a respiratory Signal Quality Index (SQI) which helps to identify moments when the extracted signals are of poor quality due to malfunctioning of the sensor. We adapted the goodness metric concept introduced by Kumar *et al.* (2015), which analysed the power level of physiological signs in frequency ranges of interest (e.g. breathing rate: 0.1Hz to 0.85Hz used in Pereira *et al.*, 2015) to assess the quality of extracted

signals. Both contact and non-contact respiration sensors are known to be influenced by motion artefacts. For example, the accuracy in measuring the respiration pattern is affected by physical activity. In the case of the chest-belt sensor, the movement of a person (e.g. while walking) may lead to changes in the belt position and hence to its tension. We extend the goodness metric proposed in Kumar *et al.* (2015) by dividing the band of interest by the total energy within a half of the sampling frequency (e.g. here 128Hz from 256Hz) to satisfy a statistical probability condition. As a respiratory SQI, we define the respiration-related goodness probability P_r as

$$P_r(\hat{f}_{\min} \leq f \leq \hat{f}_{\max}) \cong \frac{\int_{\hat{f}_{\min}}^{\hat{f}_{\max}} S_V(f) df}{\int_{total} S_V(f) df} \quad (5.8)$$

where $0 \leq P_r \leq 1$, S_V is the power spectral density in Equation (5.7), and \hat{f}_{\min} , \hat{f}_{\max} are the lower (e.g. 0.1Hz) and upper (e.g. 0.85Hz) boundaries of the expected breathing rate, respectively. Figure 5.7b shows the concept and the area between the dotted lines as the numerator of Equation (5.8). This is an equivalent form to the relative power Signal Quality Index (pSQI) which has been used as a measure of cardiac pulse signal quality (Clifford et al., 2012; Elgendi, 2016).

5.5. Results

5.5.1. Nostril-region Tracking Performance

We assessed the performance of each tracking method - Mode Seeking (Cheng, 1995), Sparse Representation (Mei & Ling, 2011), Median Flow (Kalal *et al.*, 2010) and the proposed ***Thermal Gradient Flow*** which is based on ***Optimal Quantisation*** - by computing the percentage of frames which were successfully tracked over all the trials as in Chapter 4. Before assessing the tracking performance, we briefly explore the level of motion artefacts across datasets collected from *Situation 1 (ThermalBreath I: controlled respiration in environments with thermal dynamic changes)*, *Situation 2 (ROITracking: unconstrained respiration during sedentary activity)* and *Situation 3 (ThermalBreath II: unconstrained respiration during physical activity)*.

Figure 5.8 compares the level of motion artefacts by using a Euclidian distance (unit: pixel) between the origin of a ROI in the first frame and that in the current frame. Compared with those

from the controlled conditions, the sedentary activity and physical activity produced the higher levels of motion artefacts. Furthermore, different types of motion artefacts were found according to the type of activity: high peaks of the tracker's movement due to change of one's head direction during the sedentary activity and high fluctuations of the movement due to the oscillation of the mount during the physical activity. Figure 5.9 summarises the results for the nostril-region tracking. We discuss details of each situation below.

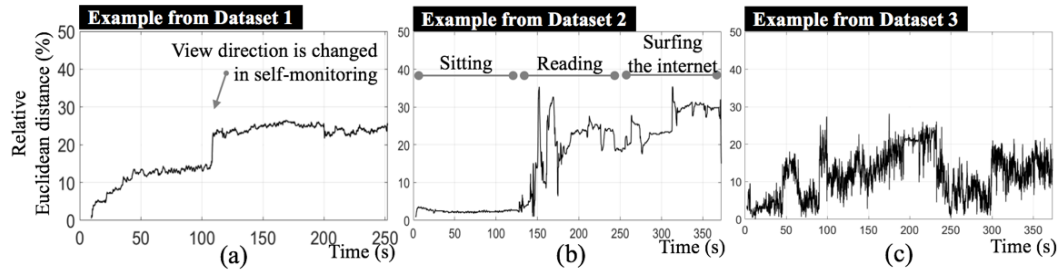


Figure 5.8. Quantified motion artefacts using the relative Euclidean distance from the origin of the nostril-ROI at the first frame: (a) example from *ThermalBreath I* (fully controlled), (b) from *ROITracking* (sedentary activity), and (c) from *ThermalBreath II* (physical activity).

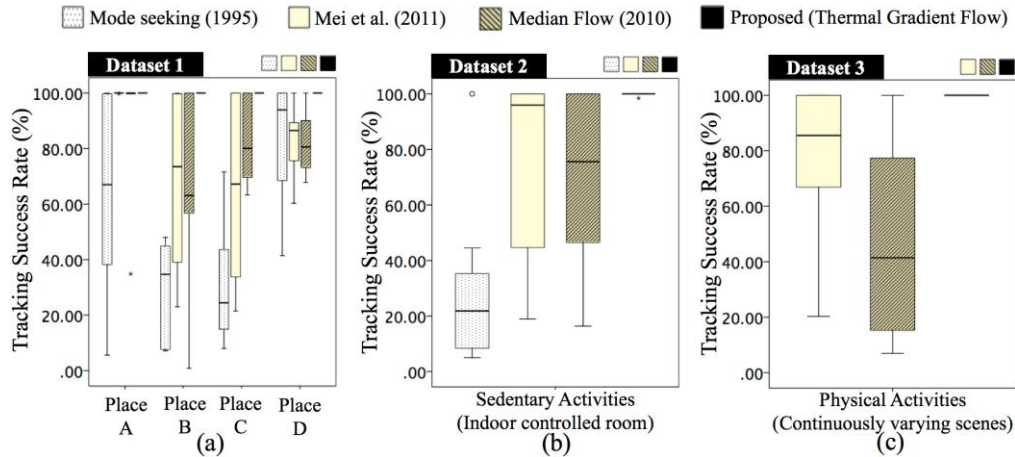


Figure 5.9. Overall results of the nostril-tracking performance of the proposed Thermal Gradient Flow (**black**) compared with existing methods: (a) *ThermalBreath I* (controlled but in non-constant temperature scenes), (b) *ROITracking* (unconstrained respiration during sedentary activity), (c) *ThermalBreath II* (unconstrained respiration during physical activity).

Controlled Respiration (ThermalBreath I): Overall results

For the data collected at *Place A* (a room, which has a *low thermal dynamic* range), while Mode Seeking performed the worst, the other three produced highly reliable results similarly as shown in Figure 5.9a (*Place A*). On the other hand, for the data collected in *Places B-D* (i.e. outside) characterized by *high thermal dynamic* ranges, the proposed approach produced significantly better results and was able to track all frames in all situations (Mode-seeking: M=47.25%, SD=33.04, Sparse Representation: M=71.31%, SD=29.37, Median Flow: M=76.35%, SD=26.16, and **our approach**: M=100.00%, SD=0.0).

As described in Section 5.4, we conducted tests of significance to examine whether the difference in the success rates across algorithms within each group (i.e. outdoor and indoor settings) was statistically significant. As some sets of the data were skewed, we used the nonparametric Friedman method. In the outdoor group (*Place B-D*) which has high thermal dynamics, there was a strong significant effect of the type of algorithm on the success rate ($\chi^2(3)=18.822$, $p<0.001$). The post hoc Wilcoxon signed ranks test (with $\alpha=0.0083 \approx 0.05/k$ where $k=6$) showed significant differences of our performance from others (with Mode-Seeking: $p=0.001$; Sparse Representation: $p=0.003$; Median Flow: $p=0.005$). On the other hand, there was no significant differences between the performance of the tracking algorithms for the indoor condition ($\chi^2(3)=7.0$, $p=0.072$). This result is consistent with findings from the literature which reports existing ROI tracking methods can perform reliably in indoor constrained situations where motion of participants and room temperature are controlled (Pereira *et al.*, 2015), indicating state-of-the-art methods were reliably implemented for our experimental evaluation.

Unconstrained Respiration (ROITracking, ThermalBreath II): Overall results

For data collected while the persons were on the move, our approach outperformed all the other methods (**Situation 2 (ROITracking) in Figure 5.9b**: Mode-seeking: M=28.47%, SD=28.37, Sparse Representation: M=76.70%, SD=31.58, Median Flow: M=67.95%, SD=34.65, and **Proposed**: M=99.84%, SD=0.49; **Situation 3 (ThermalBreath II) in Figure 5.9c**: Sparse Representation: M=78.21%, SD=26.86, Median Flow: M=48.19%, SD=35.02, and **Proposed**: M=100.0%, SD=0.0). Considering the low performance of Mode-Seeking found in the first and second datasets, we excluded this method in the comparisons for *ThermalBreath II*.

Friedman's ANOVA test results showed strong significant effects of algorithm types on success rate results with $\chi^2(3)=22.367$ ($p<0.001$) for *ROITracking* and $\chi^2(2)=17.698$ ($p<0.001$) for *ThermalBreath II*. The pairwise post-hoc analysis confirmed significantly higher performance of our proposed method than the existing state-of-the-arts in the situation where both high thermal dynamic ranges and motion artefacts existed (*ThermalBreath II*) (Median Flow: $Z=-3.18$, $p=0.001$; Sparse Representation: $Z=-2.803$, $p=0.005$, here $\alpha=0.05/3$). On the other hand, no significant differences were found between our proposed approach and the advanced tracking methods in the case of motion artefacts only (*ROITracking*) (compared with Median Flow: $Z=-2.201$, $p=0.028$; Sparse Representation: $Z=-2.023$, $p=0.043$, i.e. higher than $\alpha=0.0083\approx 0.05/k$ where $k=6$). Once again, we found that Mode-Seeking performed significantly worse than the three advanced algorithms (Sparse Representation: $p=0.008$, Median Flow: $p=0.008$, Thermal Gradient Flow: $p=0.008$). By all accounts, our method appears to be the most robust in high dynamic temperature range scenes.

5.5.2. Respiration Tracking Performance

Figure 5.10 compares Thermal Voxel-based respiration estimation method with the temperature averaging approach to extracting respiratory signals. Figure 5.10a, 5. 10c, 5. 10e show, in the time domain, the respiratory signals extracted from both methods and the ground truth (belt-sensor) for a subject randomly chosen as an example: (a) for subject S5 from Situation 1 (*ThermalBreath I*), (c) for subject S10 from Situation 2 (*ROITracking*), and (e) for subject S7 from Situation 3 (*ThermalBreath II*).

The respiration-related goodness probability P_r (see Equation (5.8)) from each time-window was computed and the overall data from every subject is described in the histogram chart in Figure 5.10b, 5.10d, 5.10f. In Situation 1 (*ThermalBreath I*), the P_r of each measurement shows similar distributions for the three methods (the **proposed** Thermal Voxel-based: $M=0.9895$, $SD=0.0103$, average-based: $M=0.9815$, $SD=0.0181$, ground-truth: $M=0.9825$, $SD=0.0174$), indicating that respiration rates calculated by each measurement were generally reliable.

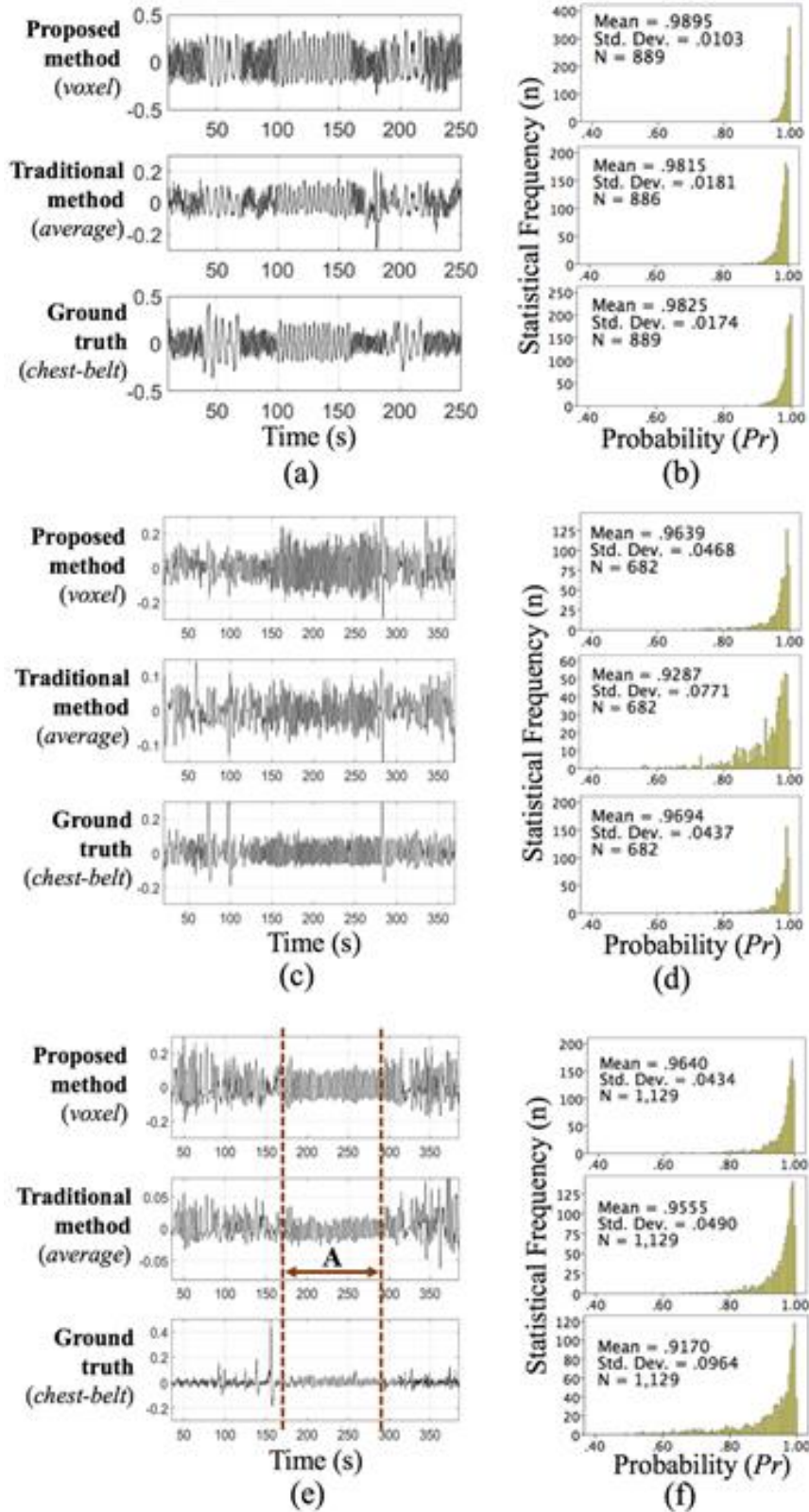


Figure 5.10. Results of respiratory signature extraction: (a,c,e) time-domain signals, (b,d,f) the respiration-related goodness metric values. The proposed Thermal Voxel-based method is more robust than the traditional averaging-based method for Situation 1 (a-b) and Situation 2 (c-d). For Situation 3 (e-f) (i.e. fully mobile contexts), the ground truth shows less reliability in the respiration tracking, except for the segment A in (e) (i.e. standing with small movement).

In Situation 2 (ROITracking) and situation 3 (*ThermalBreath II*) on the other hand, different P_r distributions were found. Overall, lower signal quality levels were observed in comparison with the controlled situations. More specifically, for Situation 2 (i.e. unconstrained sedentary activity), while the Thermal Voxel-based method and the ground truth had a similar distribution (Thermal Voxel-based: $M=0.9639$, $SD=0.0468$, ground-truth: $M=0.9694$, $SD=0.0437$), the quality of signals extracted by the traditional average-based method appeared to be more deteriorated ($M=0.9287$, $SD=0.0771$). In the case of Situation 3 (i.e. fully unconstrained physical activity), our method produced the highest signal quality while the ground truth from the belt-sensor suffered more from the physical movement (**Thermal Voxel-based: $M=0.9640$** , $SD=0.0434$, average-based: $M=0.9555$, $SD=0.0490$, **ground-truth: $M=0.9170$** , $SD=0.0964$). In particular, except for the segment involving a stationary task (i.e. labelled as A in Figure 5.10e): standing with a small movement in a dark room where there were less motion artefacts), the ground truth method produced relatively noisy patterns.

Controlled Respiration (*ThermalBreath I*): Bland-Altman and RMSE analysis

Figure 5.11 summarises the overall accuracy results of the respiratory rate estimation for Situation 1 (*ThermalBreath I*). The proposed Thermal Voxel integration method produced highly reliable performance: mean bias of 0.0882 bpm with the 95% limits of agreement being -0.7956 to 0.9721 bpm (Figure 5.11a). By contrast, the traditional averaging method produced the mean bias of 0.0755 bpm with the 95% limits of agreement being -1.8670 to 2.0179 bpm (Figure 5.11b). In addition, the Thermal Voxel-based method shows stronger correlations with the ground truth ($r=0.9987$, $p<0.001$) from the belt sensor than the traditional averaging method ($r=0.9936$, $p<0.001$) does. As summarised in Figure 5.11c, the RMSE of Thermal Voxel integration method (0.459bpm) was lower than the error of the traditional averaging method (0.993bpm).

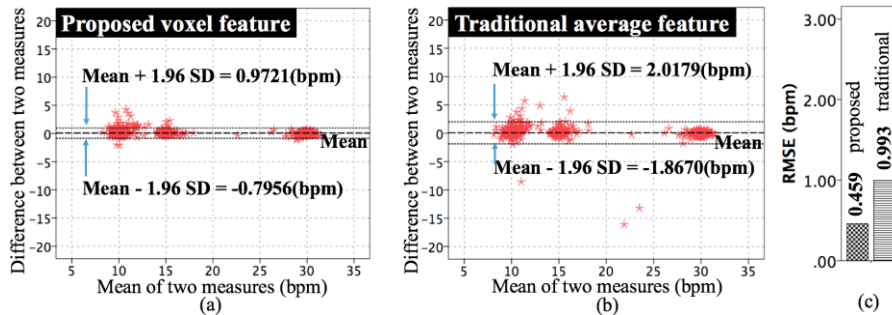


Figure 5.11. Overall results from Situation 1 (*ThermalBreath I*): Bland-Altman plots of (a) Thermal Voxel integration method, (b) the traditional averaging method, and (c) overall RMSE comparisons.

Finally, we compare the results of the methods separately over each type of environment provided in Situation 1 (*ThermalBreath I*). Figure 5.12 summarises the results of the analysis for different environments. In particular, the performance of the traditional method tends to be more affected by the different range of thermal dynamics in comparison with our method (RMSEs of the **proposed** Thermal Voxel integration method: 0.452bpm for Place A, 0.525bpm for Place B, 0.420bpm for Place C, 0.429bpm for Place D; RMSEs of the **widely-used averaging** method: 0.651bpm for Place A, 1.286bpm for Place B, 1.181bpm for Place C, 0.686bpm for Place D).

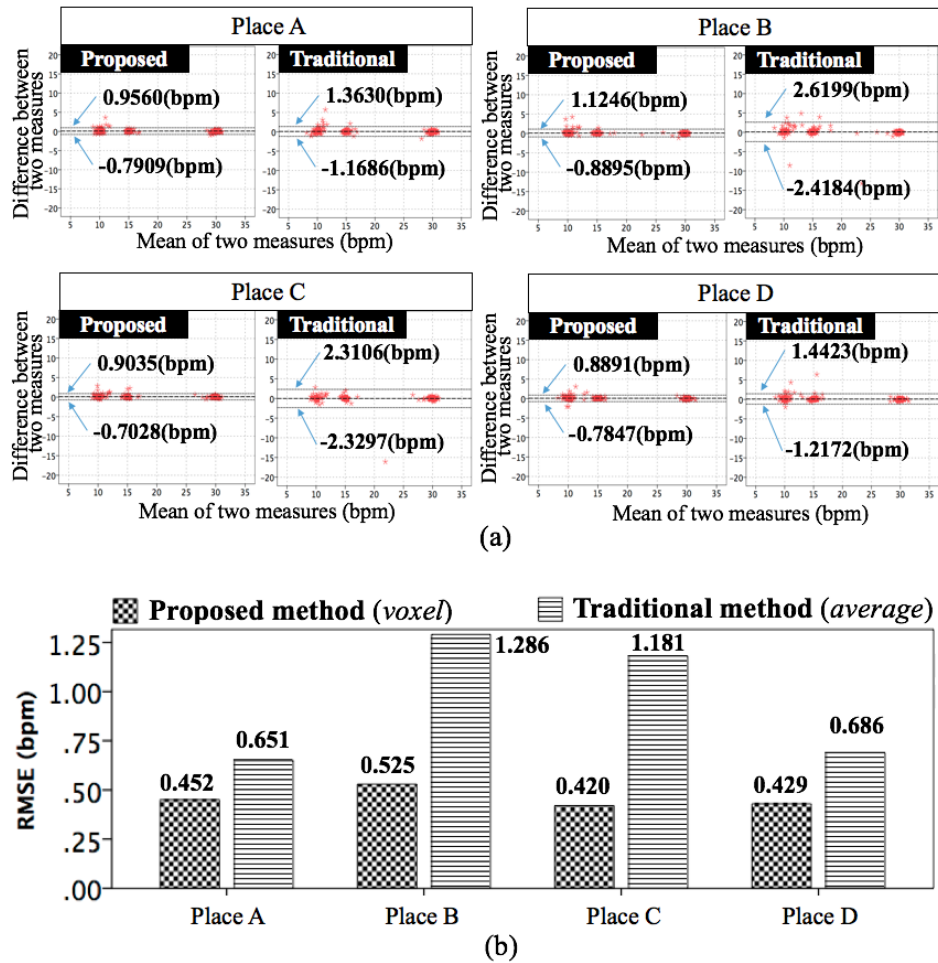


Figure 5.12. Separate results along with the different environment in Situation 1 (*ThermalBreath I*): (a) Bland-Altman plots of Thermal Voxel integration method and the traditional averaging method, and (b) RMSE comparisons for Place A – D.

Unconstrained Respiration (ROI Tracking & ThermalBreath II): Bland-Altman and RMSE analysis

Given the signal quality of the ground truth measurements for the second and third datasets were much poorer than those for Situation 1 (*ThermalBreath I*) (see Figure 5.10b, 5.10d, 5.10f), the belt sensor could not act as the ground truth. Therefore, we tested the agreement by extracting the estimated respiration rates of which pSQI is greater than or equal to the mean from the ground truth in *ThermalBreath I* (i.e. $P_r \geq 0.9825$).

For Situation 2 (*ROITracking*), Thermal Voxel integration method produced a mean bias of 0.0650 bpm with 95% limits of agreement being -1.9591 to 2.0890 bpm (Figure 5.13a), while the traditional averaging method showed a mean bias of -0.2735 bpm with 95% limits of agreement being -4.4010 to 3.8540 bpm (Figure 5.13b). Accordingly, samples derived from the proposed Thermal Voxel integration method and the traditional approach were correlated with the reference $r=0.9579$ ($p<0.001$) and $r=0.8743$ ($p<0.001$), respectively. The RMSE was also reduced more than twice from 2.11bpm (i.e. averaging method) to 1.03bpm (i.e. Thermal Voxel integration) as shown in Figure 5.13c.

For Situation 3 (*ThermalBreath II*), both methods produced less reliable results (Figure 5.14) (Thermal Voxel-based: mean bias = -0.1405 bpm, 95% limits of agreement: -4.9480 to 4.6670 bpm, $r=0.8270$ ($p<0.001$), RMSE=2.45bpm), the traditional one: mean bias=0.1921 bpm, 95% limits of agreement: -5.2682 to 4.8840 bpm, $r=0.7990$ ($p<0.001$), RMSE=2.59bpm).

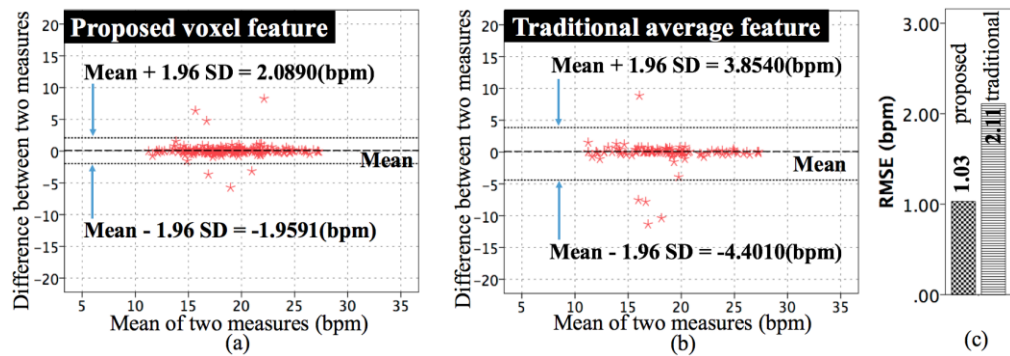


Figure 5.13. Situation 2 (*ROITracking*): Bland-Altman plots of (a) Thermal Voxel integration method, (b) the traditional averaging based estimation method, and (c) overall RMSE comparisons. The mean P_r value (pSQI) from the ground truth in Situation 1 (*ThermalBreath I*) was set as exclusion criterion ($P_r \geq 0.9825$).

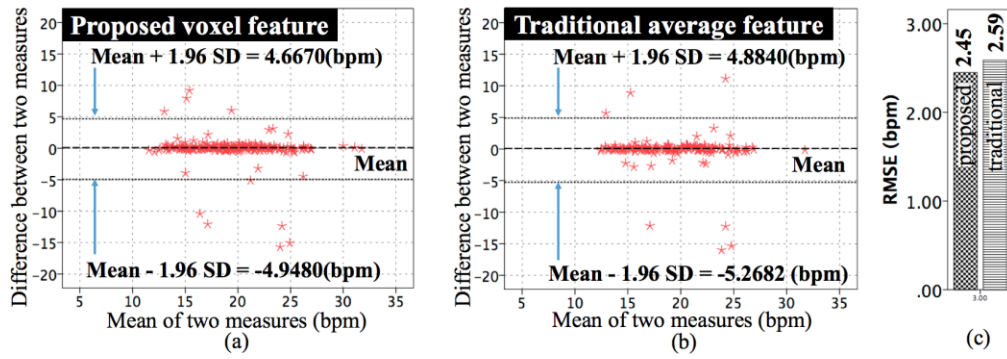


Figure 5.14. Situation 3 (*ThermalBreath II*): Bland-Altman plots of (a) Thermal Voxel integration method, (b) the traditional averaging method, and (c) overall RMSE comparisons. The mean P_r value (pSQI) from the ground truth in Situation 1 (*ThermalBreath I*) was set as exclusion criterion ($P_r \geq 0.9825$).

Finally, we demonstrate the performance of our proposed nostril-ROI and respiration tracking algorithms together through examples in [Visualisation 1](#)¹⁸ (S1 in place C from Situation 1, *ThermalBreath I*), [Visualisation 2](#) (S9 from Situation 2, *ROITracking*), and [Visualisation 3](#) (S3 from Situation 3, *ThermalBreath II*). These visualizations (available online as supplementary materials) show the results for existing state-of-the-art algorithms (Nostril-ROI tracking: Median Flow, and respiration tracking: averaging method) and for our proposed advanced approach.

5.6. Discussion

The ability of thermal imaging in extracting respiration signals in constrained indoor and stationary settings has been explored in the literature (Murthy & Pavlidis, 2006; Fei & Pavlidis, 2010; Abbas *et al.*, 2011; Lewis *et al.*, 2011; Pereira *et al.*, 2015). This chapter has aimed to enhance this power by moving into unconstrained, ubiquitous, real-world-like settings. For this, we have focused on two separate computational phases which are necessary to automate the respiration tracking process. They are: i) tracking of the nostril ROI and ii) tracking of the respiratory rate. Along with the first phase, we have proposed a novel nostril ROI tracking method named *Thermal Gradient Flow* built on the Optimal Quantisation method proposed in Chapter 4.

¹⁸ *Visualisations* are to additionally demonstrate our proposed nostril ROI and respiration tracking performance on thermal videos (and comparing ours with existing state-of-the-art algorithms). *Visualisation 1,2,3* are uploaded to the journal (Biomedical Optics Express) repository: <https://www.osapublishing.org/boe/abstract.cfm?uri=boe-8-10-4480#articleSupplMat> Details can be found from Cho *et al.* (2017c).

For the second phase, we have proposed a novel physiological signal representation technique, *Thermal Voxel Integration*.

Robust tracking of the nostril ROI in unconstrained settings

Robust nostril-region tracking is critical to monitoring respiratory signs as discussed in the works that achieved state-of-the-art performance (Lewis *et al.*, 2011; Pereira *et al.*, 2015). To the best of our knowledge, Pereira *et al.* (2015) achieved the highest performance in tracking the nostril region by adopting Mei *et al.*'s Sparse Representation method (Mei & Ling, 2011) which is one of the most advanced motion tracking algorithms. However, this was achieved in highly constrained laboratory experiments in terms of environmental temperature and motion artefacts. Outside the laboratory however, key challenges in tracking the nostril, such as noise amplified by motions and highly varying thermal dynamic ranges, have not been addressed. To overcome these challenges, we have proposed the *Thermal Gradient Flow* algorithm (proposed for dealing with motion artefacts) which is built on the top of the *Optimal Quantisation* proposed in Chapter 4 (for handling the environmental thermal change issue). We have thoroughly explored the performance of the proposed approach by comparing its performance with state-of-the-art nostril tracking algorithms.

From systematic evaluation environments to unconstrained real-world settings, the proposed method achieved new state-of-the-art performance. In particular, we found the environment with high thermal dynamic ranges greatly (negatively) influences the performance of the existing state-of-the-art methods even when participants' motion was controlled in Situation 1 (*ThermalBreath I* dataset). On the other hand, our approach was less affected by environments with changes in ambient temperature. This is an important finding helping to bring thermal imaging into more real-world-like situations.

In scenes with low thermal dynamic ranges in Situation 1 on the other hand, both the recent advanced methods (Mei & Ling, 2011), Median Flow (Kalal *et al.*, 2010) and our Thermal Gradient Flow similarly produced almost perfect results. Similarly, in the case of sedentary-motion scenario (i.e. Situation 2: *ROITracking* dataset), the three methods performed significantly better than the earlier approach (Mode-Seeking) without statistically significant differences between each other, despite the fact that our approach produced the highest performance. All in all, the results indicate benchmark algorithms were properly implemented.

In the case of tracking the nostril in fully unconstrained physical activity settings (i.e. Situation 3: *ThermalBreath II* dataset), our approach performed significantly better than all other methods, showing its robustness in the nostril-region tracking to challenges present in everyday settings, as shown in Figure 5.9.

Importance of physiological signal representation in achieving reliable respiration monitoring performance

In the body of work exploring the use of thermal imaging as a physiological sensing channel including respiration tracking, the step to represent thermal signals has by and large been overlooked by adopting the simple averaging method (Murthy & Pavlidis, 2006; Fei & Pavlidis, 2010; Abbas *et al.*, 2011; Lewis *et al.*, 2011; Pereira *et al.*, 2015). As discussed earlier, averaging tends to ignore fairly small but important changes in an array of temperature elements. Low-resolution thermal imaging makes this situation worse, losing the informative thermal variation. Hence, we have focused on newly representing thermal signatures and, accordingly, proposed the Thermal Voxel Integration method.

Through the second phase of this work, we aimed to compare our Thermal Voxel based approach with the temperature averaging method commonly used for extracting respiration patterns (Murthy & Pavlidis, 2006; Fei & Pavlidis, 2010; Abbas *et al.*, 2011; Lewis *et al.*, 2011; Pereira *et al.*, 2015). The comparison was run on the same nostril-ROI sequences which were completely tracked by our Optimal Quantisation based Thermal Gradient Flow. The results show that the Thermal Voxel integration method outperformed traditional averaging approaches.

This new approach is capable of improving accuracy, even in controlled situations designed to simulate guided respiratory rates (Situation 1: *ThermalBreath I*). It is noteworthy that the traditional temperature averaging method tends to be negatively affected by environments with a wider range of ambient temperature variances, for example, Place B and C in Figure 5.5 and 5.12, while our Thermal Voxel integration method is relatively less susceptible to these changes. Under motion artefacts (Situation 2 and 3: *ROITracking* and *ThermalBreath II* datasets), both methods showed relatively low agreement with the chest-belt respiration sensor; however, the quality of the reference signal was also low. In particular, the quality of the reference data from participants who were involved in physical activities was worst (pSQI=0.917), while our proposed method still showed high quality data (pSQI=0.964). This indicates that the breathing belt-type sensor which has been used as a gold standard breathing measurement is susceptible to

motion artefacts. This could be explained by physical activities inducing loose fitting of the sensor on one's upper body.

Limitations, future opportunities and connection to next chapters

There are still places for improvements. First, the conducted experiments do not cover all the possible scenarios. These include extreme cases where there are sudden transitions between scenes with different levels of ambient temperature (e.g. leaving a heated building in winter) or under varying levels of humidity (e.g. swimming pool, sauna) which could influence thermal patterns (Lloyd, 2013).

Secondly, the low-cost thermal camera characteristics may have contributed to some of the calculation errors. For example, thermal data collected from the system includes sporadically a few extreme values (e.g. over 100°C or lower than -30°C), wrongly calculated due to lens-inducing errors (e.g. lens focusing, misted edges of lens), which could possibly degrade the performance in respiration tracking. Fortunately, the statistical-outlier reduction used in the Optimal Quantisation technique (proposed in Chapter 4) is capable of minimising negative effects of the issue. Other than this, the resolution-related aspects are less likely to be fundamentally improved by algorithmic approaches; for instance, a one or two successive thermal image frames are sometimes blurred due to the low-quality issue, leading to temporary tracking errors. This is further discussed in Chapter 6.

Similarly, the performance of our approach may be affected by extremely fast breathing rate beyond the range of breathing rates observed in our experiments. This is due to the fact that the temporal resolution of the low-cost thermographic system we use depends on software schedulers of a mobile operating system (i.e. Android in our case) which leads to producing unsteady sampling rates. As discussed in the literature (Kumar *et al.*, 2015), higher sampling rate can directly improve the precision of respiratory indices. Given this, we expect that the better the spatial resolution of the data, the more enhanced the signals will be. All in all, it is important to note that our approach has achieved the state-of-the-art performance **even with the use of low-cost, low-resolution systems** and we can expect further enhancements in the respiration tracking performance in the case where we use higher resolution thermal imaging, for example those used in the literature (Murthy & Pavlidis, 2006; Fei & Pavlidis, 2010; Abbas *et al.*, 2011; Lewis *et al.*, 2011; Pereira *et al.*, 2015).

Finally, despite our focus on the extraction of respiratory rates in the second phase of this work, the underlying advantage of our proposed method is not just about strong agreements with the reference sensor in this respiratory index. As shown in Figure 5.10a, 5.10c, 5.10e, tracked signals contains rich, continuous information of breathing¹⁹, which is far more informative for inferring one's mental and physical condition than the rate itself. Given this, we will explore how to utilise the information and capture stress-related cues in Chapter 8.

5.7. Chapter Summary

In this chapter, we have proposed a novel respiration tracking method which reliably tracks the position of the nostril on thermal images and extracts respiration patterns from temperatures inside the nostril. To build a mobile thermal imaging-based practical respiration sensor, we have identified major challenges: difficulty in tracking the nostril and low respiratory signal quality from mobile thermal imaging in unconstrained situations. To overcome the identified issues, we have proposed a novel thermal gradient map-based ROI tracking technique, *Thermal Gradient Flow* which builds on the Optimal Quantisation proposed in Chapter 4. Furthermore, we have proposed a novel *Thermal Voxel*-based integration method to strengthen the respiratory signal quality.

With datasets collected from the conducted experiments, we have evaluated the performance of our methods in tracking the nostril region and the respiratory rate. We have achieved state-of-the-art performance (e.g. strong correlation with ground truth respiration-belt sensing data in controlled situations, $r=0.9987$. Note that the prior art performance is $r=0.974$ in Pereira *et al.* 2015). We have also demonstrated that this new nostril-tracking strategy performs significantly better in scenes with high thermal dynamic ranges compared to the three state-of-the-art algorithms.

¹⁹ An example of rich, continuous breathing information tracked by the proposed method is demonstrated in a news article which featured this work: <https://phys.org/news/2017-09-software-mobile-phone-accessory.html>

Chapter 6

Robust Tracking of Stress-induced Nasal Thermal Variability

This chapter aims to build computational methods for improving the capture of vasoconstriction/dilation related cardiovascular thermal signatures from the nasal area, whose responses to mental stressors have been explored in the literature. Key challenges which have emerged in the literature are: i) high-levels of constraints given for tracking the nose tip area and extracting temperature (e.g. control of room temperature and a person's motion); and ii) the limited capability of the widely-used single metric, that is, temperature difference between two moments (e.g. stress-induction and resting sessions), which is susceptible to noise. To address this, we focus on developing a technique that can continuously and automatically monitor the temperature of the nose tip, in turn helping minimise such constraints. We also introduce a richer set of metrics addressing limitations of the use of the single metric.

Section 6.1 discusses the background and key challenges. Section 6.2 proposes our new methods and Section 6.3 describes the conducted experimental studies and collected datasets. Evaluation methods and results are described in Sections 6.4 and 6.5, which are followed by a discussion in Section 6.6, and a chapter summary in Section 6.7.

6.1. Background and Research Question

Physiological signal sensing plays an important role in monitoring a person's mental stress. As discussed in Section 2.3, it has been shown that thermal imaging has potentials to sense vasoconstriction and vasodilation patterns from the skin (on vessels). Such patterns that cause blood flow to drop or increase are influenced not only by ambient temperatures (e.g. local cooling or warming) (Pergola *et al.*, 1993), but also mental stress (Elam & Wallin, 1987). In particular,

earlier works using thermal imaging have documented vasoconstriction-related temperature drop of the nose tip tissues is one of the key cardiovascular events which happen as a reaction to mental stressors (Veltman & Vos, 2005; Or & Duffy, 2007; Engert *et al.*, 2014; Ioannou *et al.*, 2014). In order to observe the stress-related signature, the literature has mostly focused on capturing the direction in thermal variations of the nose tip.

During data capturing, a majority of studies have drawn upon visual inspections and also imposed motion constraints (using a chinrest) (Veltman & Vos, 2005, Engert *et al.*, 2014; Salazar-López *et al.*, 2015). A few studies have used motion (ROI) tracking algorithms; however, participants have still been required to keep their head still (Abdelrahman *et al.*, 2017). As discussed in previous chapters, this is one of the main challenges, keeping thermal imaging from being used in unconstrained, real-world-like settings.

The use of ROI tracking methods proposed in Chapters 4 and 5 might help to address this challenge. However, the nose tip area is an area which is difficult to track on thermal images. For instance, Figure 6.1 shows examples of thermal images of a person's face which we collected from an office, and the nose tip area which is selected as a ROI. The shape of the local facial skin region is often blurred and not consistent on thermal images. This does not provide a sufficient number of key facial features which are required for ROI tracking. This is due to the homoeothermic metabolism and relatively low thermal conductivity of a person's skin (here, the nose tip) keeping a very narrow range of temperature distributions on it. In the case of the nose tip, the conversion of raw thermal data to a thermal-gradient magnitude map proposed in Chapter 5 (Equation (5.2)) is less likely to help obtain strong feature points from the area, as shown in Figure 6.2 right (there is no facial feature inside the bounding box on the nose tip). This leads to our research question: *Can we continuously monitor stress-induced vasoconstriction/dilation patterns from the nose tip in unconstrained settings?*

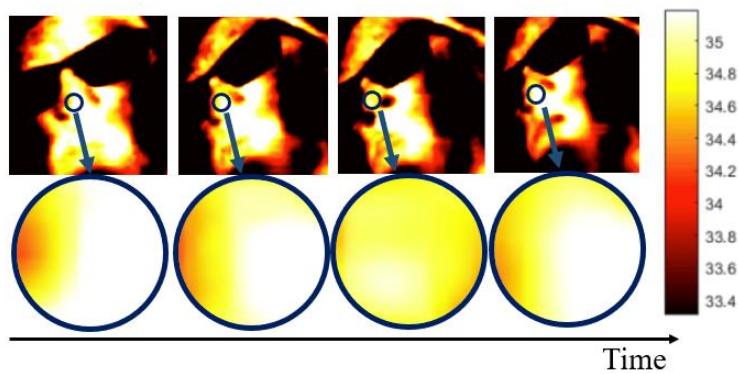


Figure 6.1. Examples of thermal images of a person's face and the nose tip selected as a ROI: the ROI from each thermal image has different temperature distribution, not providing common point features or patterns across thermal images which are required for the ROI tracking.

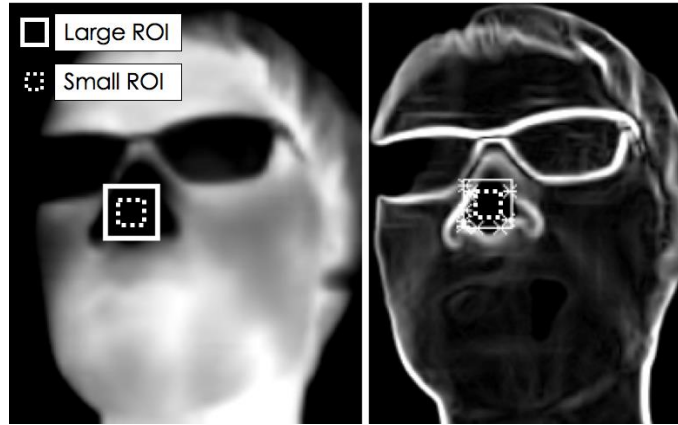


Figure 6.2. Selection of a larger ROI that includes the nose tip (small ROI) to obtain strong facial point features to enable reliable ROI tracking: **(left)** raw thermal image, **(right)** thermal gradient image. White stars (*) represent facial point features.

Another unsolved challenge is how to enrich the quality of information derived from vasoconstriction/dilation-induced temperature changes. In the literature, *temperature difference* between two time points (positive/negative direction and its amplitude in temperature change, e.g. average of -0.56°C after being exposed to stressors in Or & Duffy, 2007) have been dominantly used as a single metric. However, a person's skin temperature is influenced by physiological activities (e.g. vasoconstriction and vasodilation) as well as external factors such as environmental temperature. Hence, the single metric could be sensitive to other factors beyond the stress-induced physiological reaction. Indeed, incongruent results have been reported by studies where the metric was mainly used (Veltman & Vos, 2005 and Engert *et al.*, 2014). This indicates there is a need for building a richer set of metrics to compensate for the limited capability of a single metric, which is likely to lose important physiological information of one's mental stress. This may explain another reason why thermal imaging-based vasoconstriction and vasodilation monitoring has not yet been used in real-world applications. This leads to a follow-up research question: *Can we build a rich set of metrics to quantify variations in the patterns?*

In the next section, we propose computational methods to address the challenges mentioned above. Firstly, we introduce a technique to continuously track the nose tip temperature from peripheral nasal areas in less constrained settings. Following this, we propose a set of metrics to capture richer information from the nose tip temperature. This helps capture thermal variability rather than only the thermal directionality on which the existing metric is based (Genno *et al.*, 1997; Veltman & Vos, 2005; Or & Duffy, 2007; Engert *et al.*, 2014; Abdelrahman *et al.*, 2017; Di Giacinto *et al.*, 2014).

6.2. Proposed Computational Methods

In this section, we propose a tracking technique that builds upon the Optimal Quantisation (Chapter 4) and the Thermal Gradient Flow (Chapter 5). We then introduce a set of novel metrics.

6.2.1. Continuous Tracking of Nasal Temperature

Given the blurred and inconsistent shape of the nose tip (skin) on a thermal image (Figure 6.1 and Figure 6.2), this section proposes to select a larger ROI including the nose tip where more distinct morphological shape can be obtained for the tracking, as visualised in Figure 6.2 right, rather than the nose tip area itself (Or & Duffy, 2007; Abdelrahman *et al.*, 2017). The selection of a larger ROI enables our Optimal Quantisation-enabled Thermal Gradient Flow (proposed in Chapter 5) to continuously track the area and estimate the nose tip temperature. Figure 6.3 illustrates a series of computations we propose, which contain i) the selection, and tracking, of a larger ROI, ii) computation of spatial average of temperatures within the ROI, iii) reduction of outliers with a sliding window and iv) low pass filtering.

Following the literature (Or & Duffy, 2007; Engert *et al.*, 2014; Abdelrahman *et al.*, 2017), we compute the spatial average of temperatures in the ROI from every single frame to obtain a one-dimensional time series of thermal data, as shown in Figure 6.3 (Step II). Even if we achieve good performance in tracking ROIs, temporary errors are likely to occur due to blurred frames (e.g. just one or two frames) to which the low-resolution, low-cost thermal imaging leads (e.g. lens-inducing errors, calibration errors produced by a thermal camera). Hence, the third step is to remove such outliers related to temporary tracking errors as shown in Figure 6.3 (Step III). This can be done by excluding values beyond a range computed from Tukey's hinge ($g=1.5$) (Tukey, 1977), which has been widely used in outlier rejection processes. The range can be expressed by

$$[Q_1 - g(Q_3 - Q_1), Q_3 + g(Q_3 - Q_1)] \quad (6.1)$$

where Q_1 and Q_3 are the first and the third quartiles from temperature distribution and g is the Tukey's constant. To compute the range and remove the outliers, we use a sliding window and set the length of the window to one third of the total length of data.

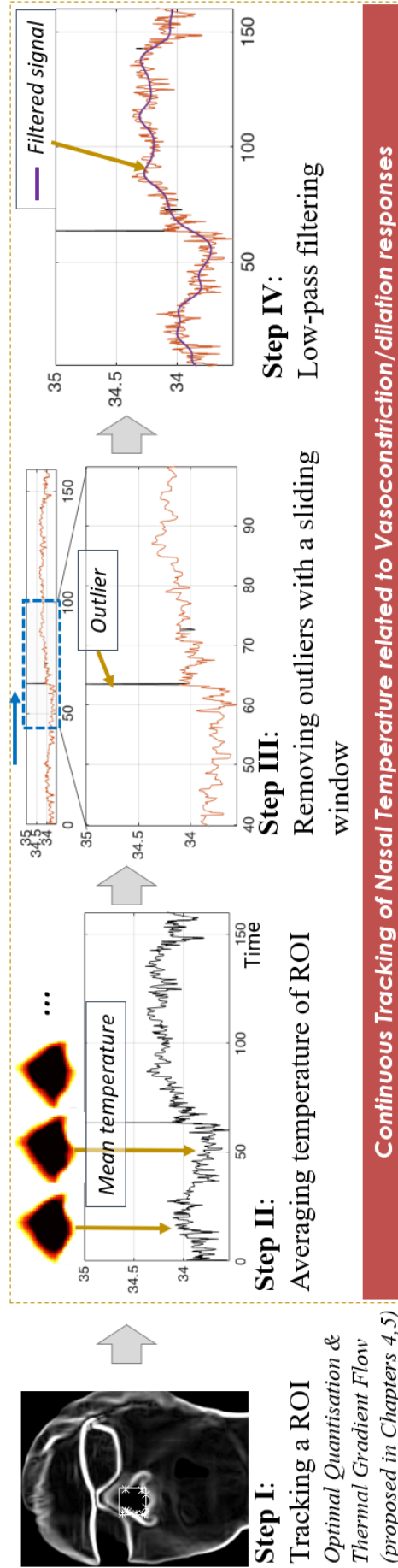


Figure 6.3. The proposed method for tracking of the nose tip temperature variances: Step I) selecting a larger ROI and tracking, Step II) computing average of temperatures within the ROI, Step III) removing outliers with a sliding window, Step IV) removing high frequency components.

Now, a remaining step is to minimise a person's respiration artefacts as i) we choose a larger ROI which includes part of nostrils and ii) breathing does also cause thermal changes in the area close to the nose tip. Figure 6.4a shows examples of thermal images taken from participants during the data collection study (introduced in the next section). From our observations, we found that respiration influenced the nasal tip temperature measurement in some cases. For instance, in Figure 6.4b, sequential thermal images of a person's nose tip surface show that inhaled air influences the nose tip temperature. This indicates a need to remove such effects from the ROI to achieve reliable measures. This is necessary despite the fact breathing information itself could be one of the significant indicators of mental stress (which is thoroughly discussed in Chapter 8). For this, we propose to use a low-pass filter with a cut-off frequency lower than the normal range of breathing rates of healthy people, for example the range [0.1-0.85Hz] considered in Chapter 5. As a thermal directional change is a relatively slow physiological event (Kuraoka & Nakamura, 2011), we set this to 0.08Hz which is lower than the low boundary. For the implementation, we used a zero-phase filtering (seventh-order, Butterworth) to avoid a phase-shifted result. Figure 6.3 (Step IV) shows the continuously tracked signal representing temperature variations of the nose tip.

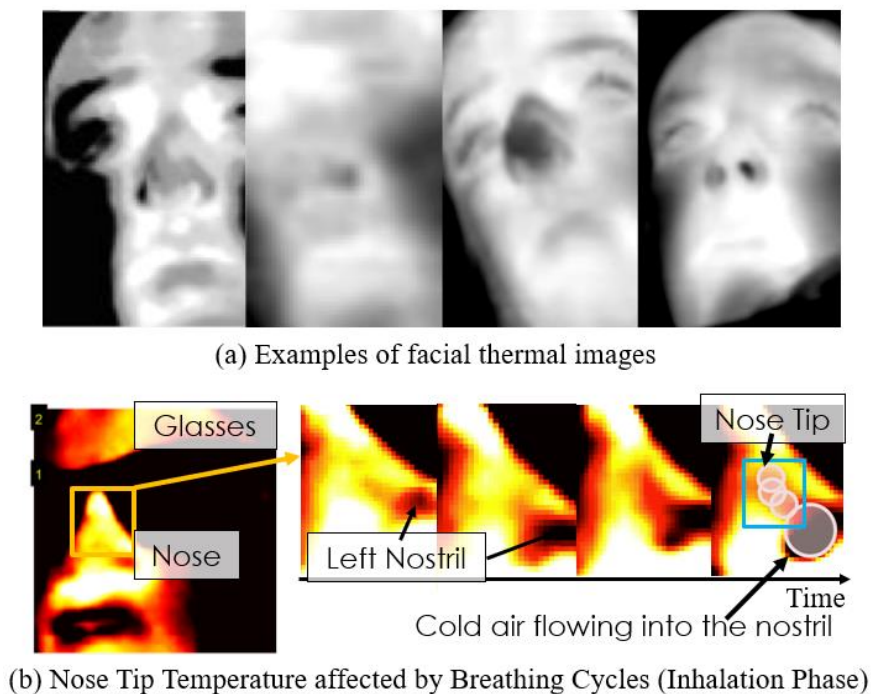


Figure 6.4. A person's respiratory activity influences the nasal tip temperature: (a) examples of thermal images (view angles were not constrained), (b) the nasal temperature changes during inhalation (yellow: warmer, red: moderate, black: colder).

6.2.2. Metrics for Quantification of Nasal Thermal Variability

This section aims to develop a possible set of metrics that can help capture thermal variability of a tracked nasal thermal variance sequence including directional information. For this, this chapter introduces a further post-processing step whose key components are i) down-sampling and ii) feature scaling, as shown in Figure 6.5. Down-sampling is to address the unsteady sampling rate of the thermal camera ($\sim 8.7\text{Hz}$) and then compute successive temperature differences sampled at regular temporal points. Here, we simply use a linear interpolation to down-sample (1Hz) the sequence. Feature scaling (Figure 6.5b) is to minimize the effect of difference of nasal temperatures across participants and sessions (called individual physiological difference, Hernandez *et al.*, 2011). The two components are of help whilst exploring thermal (temporal) variability from a single measurement. For the scaling, we used one of the widely used methods: min-max feature scaling. This considers the entire sequence collected from a session for each participant and can be expressed by

$$m(k) = \frac{x(k) - x_{\min}}{x_{\max} - x_{\min}} \quad (6.2)$$

where $k \in [0, n-1]$, $x(k)$ is the resampled sequence. With the aim to capture thermal variability, we label the resampled outcome as **thermal variability sequence** and the feature-scaled outcome as **feature-scaled thermal variability sequence** from which we produce a richer set of metrics.

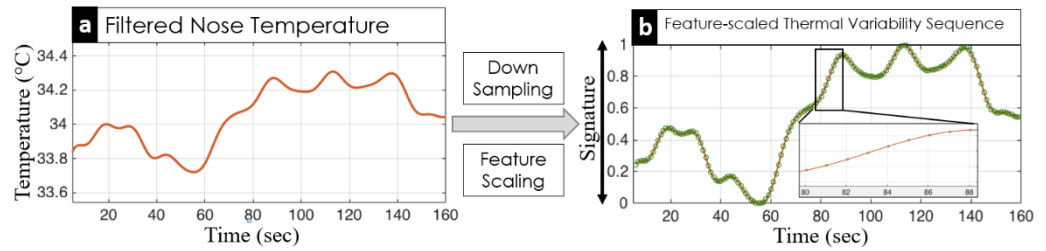


Figure 6.5. Post-processing (down sampling and feature scaling using the min-max approach) for the extraction of one-dimensional feature-scaled thermal variability sequence from a person's nose tip.

Table 6.1 describes the set of metrics we propose to represent thermal variability. The first metric is derived from the literature (Genno *et al.*, 1997; Or & Duffy, 2007; Engert *et al.*, 2014; Abdelrahman *et al.*, 2017), i.e.,

- TD (**T**emperature **D**ifference between two points, here we select temperatures from the start and the end).

Inspired by HRV (Heart Rate Variability) metrics (Pagani *et al.*, 1986; Camm *et al.*, 1996; Hjortskov *et al.*, 2004; Williamon *et al.*, 2013; Shaffer & Ginsberg, 2017), the second and third metrics are to capture variances:

- SDSTV (**S**tandard **D**eviation of **S**uccessive differences of **T**hermal **V**ariability sequence, i.e. the difference between each interval on the down-sampled data),
- SDTV (**S**tandard **D**eviation of **T**hermal **V**ariability sequence).

Additionally, the fourth metric is to capture a global thermal directional change by using the slope from a linear polynomial fitting (used in Engert *et al.* 2014):

- STV (**S**lope of **T**hermal **V**ariability sequence).

We applied the four metrics to our proposed feature-scaled thermal variability sequence and TD and STV to non-feature-scaled sequence (explored in the literature) for the comparison purpose. This set of metrics are summarised in Table 6.1 with their supporting equations.

Table 6.1. A novel set of metrics to quantify Thermal Variability

Metrics	Definition and Equation
<i>Feature scaled thermal variability sequence, $m(k)$, $k \in [0, n-1]$</i>	
TD_F	<p>Temperature Difference between <i>feature scaled</i> data from the start and the end</p> $\Delta T_m = m(n-1) - m(0)$
$SDSTV_F$	<p>Standard Deviation of Successive differences of <i>feature scaled</i> Thermal Variability sequence</p> $s_{\hat{m}} = \sqrt{\frac{\sum (\hat{m} - \bar{\hat{m}})^2}{n-1}}, \quad \hat{m}(k) = m(k+1) - m(k)$
$SDTV_F$	<p>Standard Deviation of <i>feature scaled</i> Thermal Variability sequence</p> $s_m = \sqrt{\frac{\sum (m - \bar{m})^2}{n-1}}$
STV_F	<p>Slope of <i>feature scaled</i> Thermal Variability sequence</p> $\beta_m = \beta_1, \quad y = \beta_0 + \beta_1 m + \varepsilon \text{ (from linear polynomial fitting)}$
<i>Thermal variability sequence, $x(k)$, $k \in [0, n-1]$</i>	
TD	<p>Temperature Difference between data from the start and the end*</p> $\Delta T_x = x(n-1) - x(0)$ <p>* This is derived from the existing, widely used, metric (Genno et al., 1997; Or & Duffy, 2007; Engert et al., 2014; Abdelrahman et al., 2017).</p>
STV	<p>Slope⁺ of Thermal Variability sequence</p> $\beta_x = \beta_1, \quad y = \beta_0 + \beta_1 x + \varepsilon \text{ (from linear polynomial fitting)}$ <p>⁺ This is derived from the slope used in Engert et al. (2014).</p>

6.3. Experimental Protocols and Datasets (*NoseTracking*, *StressNose*)

The main aim of the data collection is two-fold: i) to systematically evaluate the use of a larger ROI on thermal images collected from carefully controlled situations (Dataset I: *NoseTracking*), and ii) to verify the capability of the proposed metrics in quantifying mental stress induced by sedentary stressful tasks (Dataset II: *StressNose*).

NoseTracking dataset: Relaxation on a Chin Rest

Figure 6.6 shows the experimental setup for the data collection (the image was taken by a thermal camera). The aim was to compare outcomes produced by the proposed method with a reference signal (a temperature sequence from a small ROI only containing the nose tip). Hence, participants' motion had to be carefully controlled. Following Veltman & Vos (2005), we used a chin rest on which each participant relaxed in order to maintain the position of her/his nose to remain as still as possible.

10 healthy adults (aged 22-50 years, 4 females) took part in the data collection study. The study took place in a quiet lab room with no distractions (and no room temperature control). Thermal image sequences were recorded using a low-cost thermal camera (FLIR One 2G) integrated into a smartphone which was placed in front of each participant (circa 50cm).

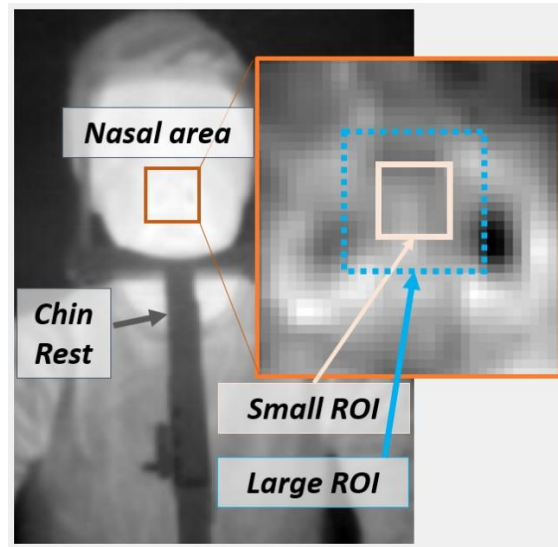


Figure 6.6. Experimental setup (the image was taken by a thermal camera) for the data collection for Dataset I: a chin rest was used to maintain the position of the nose of each participant still. The small ROI (reference) includes only the nose tip and the large ROI contains the nose tip and its surrounding area (e.g. part of the nostril).

StressNose dataset: Stress-induction task using Mathematical Serial Subtraction

In this experiment, amongst two widely used stress induction tasks discussed in Section 3.3, we used the mathematical serial subtraction task (denoted as *Math*) to induce mental stress for the stress induction. We invited 12 healthy adults (6 females) (aged 18-60 years) from the subject pool service of University College London. Each participant was given the information sheet and informed consent form prior to data acquisition. Ambient temperature and participants' movements were not controlled.

This task was divided into a resting period, experimental (*MathHard*) and control (*MathEasy*) sessions. Before starting the experimental and control sessions, participants were asked to rest for 5 minutes. The experimental condition required the participants to repeatedly subtract (mentally) a two-digit number, such as 13, from a four-digit number (e.g. 5000), for 5 minutes. In the control session, the subtracted number was set to 1, transforming the task to an easy counting-down test with the aim of inducing significantly lower stress levels (or to not induce stress). Both control and experimental sessions were counterbalanced. The task flow is summarised in Figure 6.7. Between the sessions, participants were asked to take a break, encouraging participants to fully recover (without any measurements or constraints), in order to avoid potential effects from previous sessions.

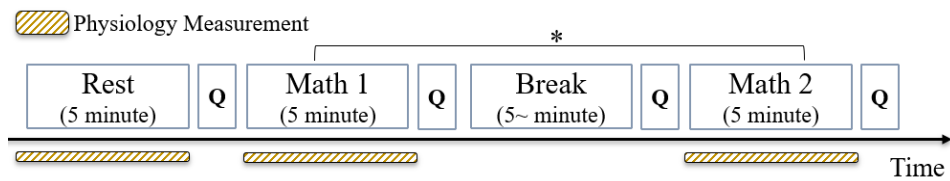


Figure 6.7. Flow chart of the study protocol (*Counterbalanced in Latin squared design).

The recording setup is shown in Figure 6.8. During the experiment, each participant answered the mental arithmetic questions using a mouse on a laptop and was thermal-video-recorded using the low-cost thermal camera (FLIR One, connected to an Android smartphone). After answering each question, participants received sound feedback which informs them whether the answer was correct or not. Before and after control and experimental sessions, participants were asked to fill in a short questionnaire (denoted as Q in Figure 6.7) to report their perceived stress level. As explained in Section 3.3, we used the Visual Analogue Scale (VAS) (Bijur et al., 2001; Lesage et al., 2012) to have participants self-report their stress scores. Only one experimenter was present in the room during the data collection but kept his distance from the participant (further than 1.5 m). All the tests were programmed and run in MATLAB (2015b, The MathWorks). The program sources are publicly available from the author’s GitHub repository²⁰.

²⁰ <https://github.com/deepneuroscience/Paced-Math-Test> (*DeepBreath* project: Cho et al., 2017a, 2017b).

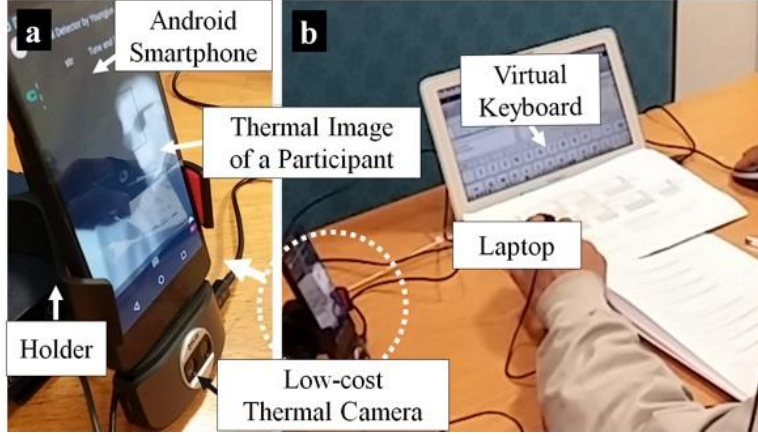


Figure 6.8. Experimental setup: a smart phone with a low-cost thermal camera was installed and running on the desk while each person participated in the tasks. The participant was not asked to constrain their movements.

6.4. Evaluation Method

Evaluating the use of a Larger ROI in tracking Nasal Thermal Variance

The first evaluation phase aims to check if nasal thermal variances extracted from the larger ROI which includes the nose tip and its surrounding areas can be a surrogate for the nose tip temperature changes. Using the first dataset *NoseTracking* where motion artefacts were fully controlled, we were able to directly compare the average temperature collected from the large ROI with that from the small ROI (only including the nose tip area which has been thoroughly explored in the literature), excluding effects of ROI tracking. We compared the Pearson correlation coefficients.

We then tested how much participants' respiratory cycled events affected nose tip temperature measurements and how much the filtering method handled the effects. For this, we used the relative power Signal Quality Index (pSQI), which is to assess the strength of physiological signals in a frequency range of interest, as a measure of quality (Clifford et al., 2012; Kumar et al., 2015; Elgendi, 2016), seen in Chapter 5. The respiratory pSQI for the raw thermal variability sequence can be expressed by:

$$P(\hat{f}_{\min} \leq f \leq \hat{f}_{\max}) \cong \frac{\int_{\hat{f}_{\min}}^{\hat{f}_{\max}} S_x(f) df}{\int_{total} S_x(f) df} \quad (6.4)$$

where $0 \leq P \leq 1$, S_x is the power spectral density of $x(k)$, and \hat{f}_{\min} , \hat{f}_{\max} are the lower and upper boundary of expected Breathing Rates (BR), respectively. Here, we set the expected BR range to [0.1Hz, 0.85Hz], as seen in Chapter 5.

Evaluating Metrics for Stress-related Thermal Variability

The aim of the second phase was to evaluate the capability of the proposed thermal variability metrics in stress assessment for less controlled settings where participants' motion was not controlled. For this, the second dataset *StressNose* was used. We investigated effects of a session type (i.e. Rest, Control: *MathEasy* and Experimental: *Mathhard*) on self-reports of mental stress levels. This was to check whether the study protocol was designed properly. Like studies using the VAS (described in Section 3.3) to assess perceived stress levels, we planned to use an one-way repeated measures analysis of variance (ANOVA) test. However, scores for Rest were not normally distributed, violating assumptions of the ANOVA. Hence, we instead used Friedman's ANOVA that can be performed on data with non-normality. For post hoc pairwise comparisons, we used the nonparametric Wilcoxon signed ranks test. Similarly, we used a one-way repeated measures ANOVA with a session type as an independent variable and each metric value as a dependent variable as collected metrics data met the assumptions of the ANOVA test (e.g. normality and sphericity).

6.5. Results

6.5.1. Comparison between Small and Large ROIs

From the first dataset with 10 participants, we collected four sets (average temperature and filtered average temperature sequences from the both ROIs) of 1000 thermal elements (10 participants x 100 temperature samples from the 1Hz resampled sequence of 100s). The size of each chosen ROI was: i) height (M=8.6 pixels, SD=2.17), width (M=8.5 pixels, SD=2.37) for Small ROI; ii) height (16.2 pixels, SD=4.39), width (M=23.7 pixels, SD=7.82) for Large ROI. Figure 6.9 shows an example of the four sets taken from one participant's thermal images (P2) during the data collection study. Here, we correlated each average temperature timeseries from a

small and a large ROI and then tested correlations between each filtered data from both ROIs. Overall, both data from the large ROI maintained a high correlation with the data from the small ROI ($r=0.999$, $p<0.001$ for both cases) as shown in scatter plots in Figure 6.10.

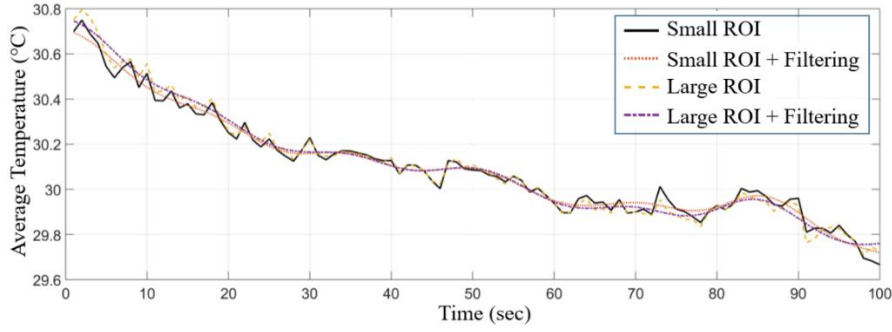


Figure 6.9. An example of extracted average temperature and (temporally) filtered temperature sequences from a small ROI and a large ROI on facial thermal images of Participant 2.

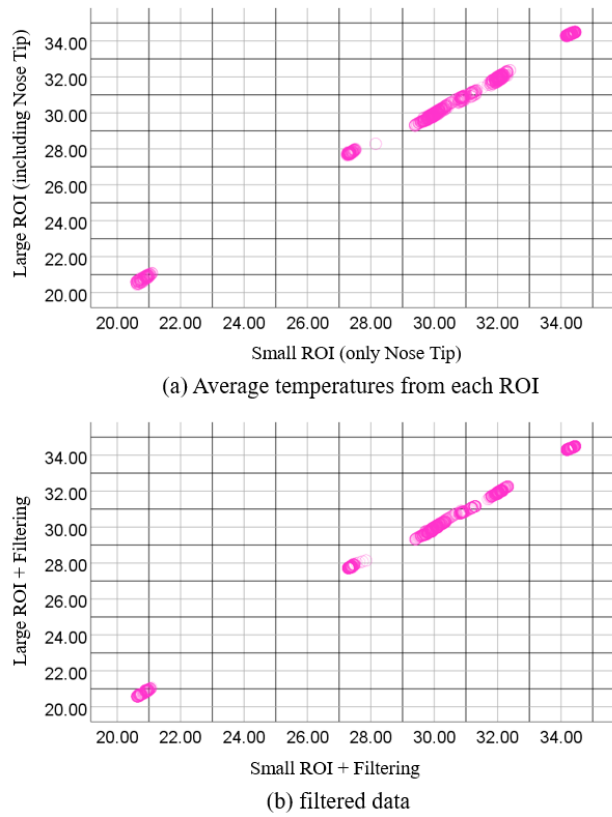


Figure 6.10. Scatter plots of (a) average temperatures and (b) (temporally) filtered average temperatures from the chosen small ROI and the large ROI (from 10 participants, $N=1000$).

As a wide range of temperatures from every participant could bring the power in achieving high correlation coefficients, we took a look at individual data and correlated them from each pair (small and large ROIs) for every participant. From this, we obtained 10 Pearson correlation coefficients from 10 participants (we correlated each pair of 100 samples for each participant). As it can be seen in Figure 6.11, the results from correlations of individual data still showed high levels of correlations between average temperature timeseries from the small and large ROIs ($M=0.926$, $SD=0.055$). Temporally filtered average temperatures, which were to be less affected by breathing artefacts, had stronger correlations between both ROIs ($M=0.987$, $SD=0.016$).

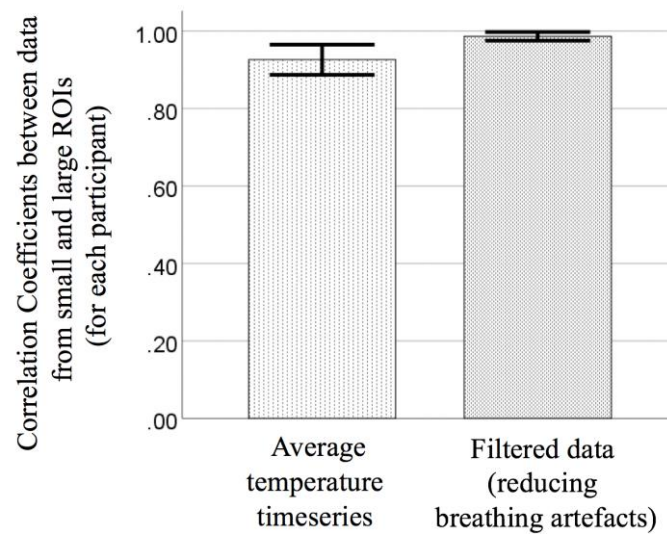


Figure 6.11. Bar plot of Pearson correlation coefficients between each pair of samples (both data from small and large ROIs) for every individual (10 participants, each coefficient was computed from each participant's samples ($N=100$)): (a) average temperature timeseries and (b) the filtered data.

As mentioned in Section 6.4, we explored levels of respiratory artefacts contained in the collected data using the relative power SQI (pSQI). Figure 6.12 demonstrates how good respiratory quality each extracted nasal temperature data has. Interestingly, average temperature time series from both ROIs (small and large) strongly engaged with respiratory cyclic events (respiratory pSQI from the small ROI: $M=0.681$, $SD=0.058$; from the large ROI: $M=0.692$, $SD=0.127$) in comparison with the filtered data ($M=0.023$, $SD=0.013$) whose pSQI was significantly different from others ($p<0.001$). This indicates that, regardless of the ROI size, temperature from the nasal area is affected by breathing, potentially leading to wrong measurements. This strongly supports a need to filter respiratory information on vasoconstriction/dilation-related nasal thermal variance data.

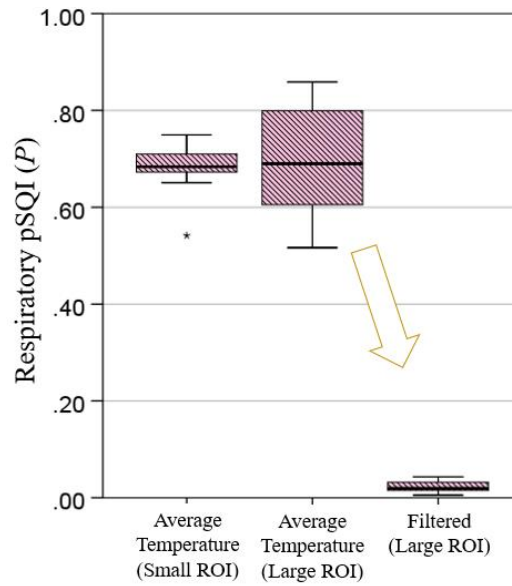


Figure 6.12. Box plots of respiratory signal quality test results about average temperatures from both small and large ROIs and the filtered average temperatures from the large ROI (the proposed signal to use) using the respiratory pSQI.

6.5.2. Effect of Stressors on Thermal Variability Metrics

First, we analysed self-reported stress scores with the aim of evaluating whether the stress-induction protocol was properly conducted. Boxplots in Figure 6.13 show the distributions of the self-reported scores across each session. Overall, it is clear that the stress elicitation procedure did produce significantly higher stress levels under the experimental condition (*MathHard*) than the ones from the other conditions (*Rest*: $M=1.46$, $SD=1.99$; *MathEasy*: $M=2.63$, $SD=1.68$; *MathHard*: $M=5.43$, $SD=2.74$), despite having one outlier from one participant (P7, *Rest*).

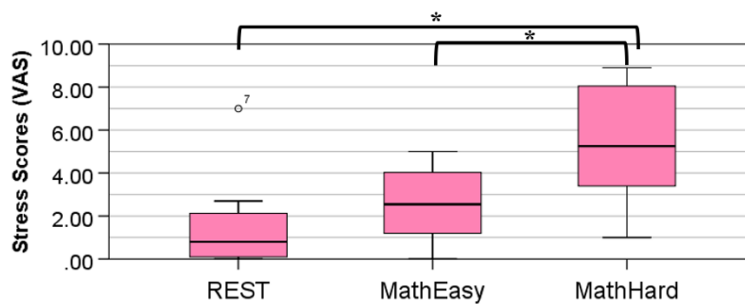


Figure 6.13. Boxplot (95% confidence interval) of the self-reported perceived mental stress scores of 12 participants across each task (*Rest*, *Math Easy*, *Math Hard*). * describes the significant difference. o⁷ (P7, *Rest*) is an statistical outlier.

Although distributions of scores collected during both arithmetic tasks were sufficiently normally distributed (MathEasy: $W=0.931$, $p=0.394$; MathHard: $W=0.927$, $p=0.347$ from the Shapiro-Wilk test), the data from the rest period were skewed ($p=0.002$). Hence, we carried out the Friedman test on the 12 participants' scores. We found significant differences between each session (including the resting period) ($\chi^2(2)=11.261$, $p=0.004$). Post-hoc Wilcoxon signed ranks test showed the experimental session (*MathHard*) elicited significantly higher stress levels than other tasks (Rest – Math Hard: $Z=-2.667$, $p=0.008$; Math Easy – Math Hard: $Z=-2.937$, $p=0.003$). On the other hand, there was no significant effect of the session type on participants' VAS scores over the resting and easy math sessions ($Z=-1.609$, $p=0.108$), indicating other components involved in tasks (e.g. using a mouse to answer questions) did not significantly affect stress levels. Hence, we could take difficult mental arithmetic as a mental stressor.

For testing the effect of a session type on the metrics in Table 6, we first examined normality of the computed data. Data from a participant (P2) collected during the rest period had extreme values making every metric value for the resting condition not normally distributed (e.g. **TD_F** for Rest: $p=0.031$ from the Shapiro-Wilk test); hence, we discarded the participant's data so that we had normally distributed data (e.g. **TD_F** for Rest without P2: $p=0.321$ from the Shapiro-Wilk test). Then, we conducted the ANOVA analysis with repeated measures. Table 6.2 provides a summary of statistical results. The results showed a significant effect of the session type on **TD_F** ($F(2,20)=10.297$, $p=0.001$, $\eta_p^2=0.507$), **STV_F** ($F(2,20)=6.089$, $p=0.009$, $\eta_p^2=0.378$), **TD** ($F(2,20)=6.847$, $p=0.005$, $\eta_p^2=0.406$), **STV** ($F(2,20)=5.78$, $p=0.01$, $\eta_p^2=0.366$) and approaching significance on **SDTV_F** ($F(2,20)=2.876$, $p=0.08$, $\eta_p^2=0.223$).

The post-hoc paired t-test with Bonferroni correction on each metric showed that only **STV_F** (our feature scaling-based metric) had a significant difference between *resting* and *MathHard* sessions ($p=0.013$) with no significant difference between MathEasy and resting sessions ($p=0.136$). This indicates that feature-scaling helps enhance the capability of thermal directionality in assessing mental stress.

Table 6.2. Significance test to assess the effect of a session type on each metrics derived from the proposed filtered sequence using a one-way repeated measures ANOVA

Source	Measure	Sum of squares	df	Mean square	F	p-value	Partial Eta squared
Task	TD_F	4.477	2	2.238	10.297	0.001	0.507
	SDSTV _F	8.88E-05	2	4.44E-05	0.628	0.544	0.059
	SDTV _F	0.005	2	0.003	2.876	0.08	0.223
	STV_F	110.135	2	55.068	6.089	0.009	0.378
	TD	4.823	2	2.412	6.847	0.005	0.406
	STV	126.792	2	63.396	5.78	0.01	0.366
Error (Task)	TD _F	4.348	20	0.217			
	SDSTV _F	0.001	20	7.07E-05			
	SDTV _F	0.018	20	0.001			
	STV _F	180.871	20	9.044			
	TD	7.044	20	0.352			
	STV	219.377	20	10.969			

6.6. Discussion

As discussed in Chapter 2, the literature has primarily focused on capturing thermal directional changes on certain facial areas in association with a person's affective states (Genno *et al.*, 1997; Pavlidis *et al.*, 2002, 2012; Gane *et al.*, 2011; Engert *et al.*, 2014). Amongst areas, the nose tip has been shown to be a key area showing a significant decrease in temperature in response to mental stressors and mental workload (Genno *et al.*, 1997; Veltman & Vos, 2005; Or & Duffy, 2007; Engert *et al.*, 2014; Abdelrahman *et al.*, 2017). This chapter has aimed to contribute to this body of work by building computational methods that can more reliably capture richer information of this physiological phenomenon.

Reliable, continuous tracking of nasal thermal variances

Throughout previous chapters, we have discussed the importance of automated ROI tracking in supporting situations beyond controlled laboratory settings where participants' mobility is highly constrained. Automatizing ROI tracking can also be a key player in innovating the extraction of nasal thermal variances as the literature has mostly relied upon motion-constrained settings (often using a chinrest) so as to enable visual inspections on thermal images to work (Veltman & Vos, 2005, Engert *et al.*, 2014; Salazar-López *et al.*, 2015). However, tracking the nose tip area on thermal images is not an easy task since its shape gets often blurred

due to many factors such as motion and breathing, which low-resolution, low-cost thermal cameras make it even worse as discussed in Section 6.1. Hence, we have addressed this by proposing the selection of a larger ROI including the nose tip and surrounding areas (e.g. part of the nostril) and conducting systematic evaluation.

Use of the ROI as a surrogate of the nose tip area has not been explored in the literature. Hence, we have investigated their relationship on the *NoseTracking* dataset which we collected in highly motion-constrained environments and found strong correlations. This is highly encouraging considering the possibility of deploying thermal imaging in real world contexts. In addition, continuously tracked signatures can be of use to capture richer information than a just discrete value such as a temperature difference. This is important as the nasal thermal variation occurs with a dynamic process of narrowing and widening of nasal peripheral vessels (vasoconstriction/dilation) under stressful conditions (Elam & Wallin, 1987; Ioannou *et al.*, 2014). Given this, we have also focused on building a set of metrics to capture richer information.

Importance of capturing multiple aspects of nasal thermal variability in assessing mental stress

The proposed tracking method allows for taking into account temporal variability of the nose tip temperature changes rather than a single discrete quantity to represent thermal directionality dominantly explored in the literature. Capturing multiple aspects of physiological variability has been shown to be important in assessing mental stress levels; for example, multiple metrics have been proposed to capture heart rate variability related to mental states (Billman, 2013; McDuff *et al.*, 2016). However, this has not yet been explored for vasoconstriction/dilation related nasal thermal variation. Therefore, we have proposed a novel set of metrics with the aim to capture complex aspects of the physiological phenomenon beyond just a temperature difference and tested their capability in quantifying mental stress on the *StressNose* dataset.

The main component we used for building metrics is feature scaling of a nasal thermal variability sequence. As we discussed in Section 6.2.2, the aim was to consider interpersonal variability in physiology which has been repeatedly reported in the literature exploring other signals (e.g. skin conductance in Hernandez *et al.*, 2011). From the experiment, we found feature scaling-applied metrics could bring more power in assessing mental stress than those to which feature scaling is not applied. The finding is interesting as it helps address limitations of earlier studies which have generally ignored individual differences in physiological patterns

(Abdelrahman *et al.*, 2017). In addition, it supports the need to track nasal thermal variances continuously. In situations where we have to collect limited samples of temperatures discretely (e.g. the start and end), it is difficult to consider individual differences.

This work also highlights the importance of avoiding collapsing complex phenomena of variability into a single metric. In our study, using the dominantly used metric *temperature difference* alone did not help discriminate the stressful event from the baseline. One possible reason for this could be the fact we did not constrain room temperatures influencing our skin temperature. The proposed set of metrics compensated for this, stressing the importance of capturing richer information of thermal variability beyond thermal directionality particularly in unconstrained environments.

Limitations, Future Opportunities and Connection to Next Chapters

Despite the above findings, there are limitations opening up future research opportunities. First, we used the spatial averaging method to summarise two dimensional thermal information on the nasal area (dominantly used in the literature). This could lead to losing important vasoconstriction/dilation induced local thermal variances. This is similar to the issue of averaging respiratory thermal variances inside the nostrils, which we addressed in Chapter 5. Hence, we further explore how to achieve better quality of spatial representation of thermal signals than averaging in Chapter 7. Second, this work drew upon hand-engineered metrics in quantifying mental stress. As hand-engineering even carefully can hardly capture all complex aspects of signals (LeCun *et al.*, 2015), we explore how to automatically learn informative features using machine learning algorithms in Chapters 8 and 9.

6.7. Chapter Summary

In this chapter, we have proposed a novel method to continuously track stress-induced thermal variability patterns on a person's nasal area. From the literature, we have identified the following issues: i) difficulty in tracking the nose tip area and its temperature continuously under motion artefacts and ii) a limitation of using a single metric in quantifying mental stress. Addressing the issues, this chapter has contributed: i) a novel tracking technique that is built on the Optimal Quantisation and Thermal Gradient Flow methods proposed in previous chapters, ii) a novel set

of metrics that can provide richer information of stress-induced cardiovascular events on the nose tip than the single metric designed to capture only the thermal directionality. We have evaluated the contributions with a series of experiments: one was conducted with 10 participants in a highly controlled situation for a systematic evaluation, and the other was conducted with 12 participants in unconstrained situations with stressful arithmetic solving tasks. The findings of this chapter is linked to Chapter 9 where we aim to build an instant stress recognition system using continuously tracked nasal thermal variability sequences.

Chapter 7

Exploring Stress-induced Cortical Thermal Signatures from the Scalp

Most work to date has explored facial temperature variation in relation to mental stress. However, as described in Section 2.4, mental stress induces changes in regional Cerebral Blood Flow (CBF) which could engage with heat propagation throughout the scalp. Hence, in this chapter we aim to explore whether stress-related thermal signatures can be captured from the scalp. The ultimate goal is to build a novel low-cost, mobile, non-contact way for observing cortical activities, which could be useful in many real-world applications (Boto *et al.*, 2018). In this chapter, we focus on exploring scalp temperature under sedentary stressful conditions. As explored in the previous chapters, use of low-cost mobile thermal imaging can support an easy set-up, particularly for near-a-person's scalp monitoring in comparison with high-cost, heavyweight thermal imaging systems.

In Section 7.1, we review cortical regions which have been shown to respond to mental stressors from brain imaging studies and discuss possible metrics for quantifying cortical thermal signatures. Section 7.2 introduces an initial investigation of cortical thermal signatures in a study we conducted within unconstrained sedentary desk activities. It reports findings and challenges which are considered in designing the protocol of a second controlled study in Section 7.3, where we investigate a representation of cortical thermal signatures and propose metrics to quantify the represented signatures. The chapter ends with a discussion and chapter summary in Section 7.4 and Section 7.5.

7.1. Background and Research Question

As discussed in Chapter 2, while the majority of studies in the literature have focused on a person's facial area to identify if certain affective states induce binary directional changes in temperature, there is evidence that other body regions could also provide such thermal signatures. *The scalp* is the region this chapter focuses on. The very initial, however direct evidence of its thermal relation to brain activities can be found from Lombard's earlier work one and half centuries ago (Lombard, 1878). Using surficial mercury thermometry, the author showed how cognitive and emotional functions were associated with an emission of heat from the brain. Although, ever since Lombard's discovery, there has been no rigorous investigations on, and attentions to, the scalp temperature in relation to emotions or affects. Indirect evidence can be found from recent neuroscientific studies on Cerebral Blood Flow (CBF) of local regions in the brain near the scalp, which engages in a cerebral heat production potentially influencing scalp temperature (Busto *et al.*, 1987; Kiyatkin, 2010; Wang *et al.*, 2014a; Li *et al.*, 2017). From this evidence, we hypothesise thermal imaging may capacitate to capture CBF activation-induced thermal signatures from certain local regions on the scalp.

Stress-activated brain areas

It has been observed for over two decades that regional changes in CBF can occur during stress response amongst other types of affective responses. From studies using fMRI, PET, and fNIRS which are used to measure CBF patterns in the brain (Taylor *et al.*, 1997; Soufer *et al.*, 1998; Wang *et al.*, 2005; Pruessner *et al.*, 2008; Dedovic *et al.*, 2009; Causse *et al.*, 2017; Rosenbaum *et al.*, 2018), one of the main regions commonly activated in response to mental stressors is the *left inferior frontal gyrus* (Taylor *et al.*, 1997; Soufer *et al.*, 1998; Pruessner *et al.*, 2008). This region is known to be involved in semantic and phonological processing (Costafreda *et al.*, 2006). Using the Neurosynth platform²¹ (Yarkoni *et al.*, 2011), Figure 7.1 illustrates the location of the region in the MNI coordinate space (with x,y,z values, D: Dorsal, V: Ventral, A: Anterior, P: Posterior, L: Left, R: Right). This visualisation is helpful to localise the ROI on a thermal image of the scalp. From the *z-axis* image (on which “*z=-12*” is written) in Figure 7.1, we can see that the region is situated near the scalp. At the same time, the location is

²¹ *Neurosynth* is a web-based platform for large-scale, automated synthesis of fMRI data (<http://neurosynth.org/>).

in the low-middle on the x -axis image (“ $x=-44$ ”), indicating the scalp frontal view would be more appropriate than the top view for the thermal camera to be able to capture the ROI.

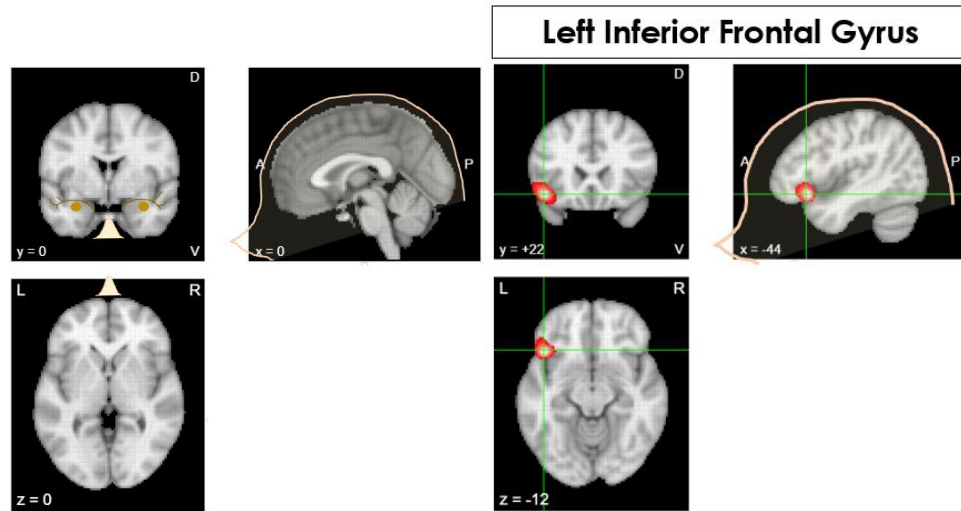


Figure 7.1. Location of the *left inferior frontal gyrus* (Red and yellow colour), the most commonly activated region under stress conditions (Taylor *et al.*, 1997; Soufer *et al.*, 1998; Pruessner *et al.*, 2008): on the brain images synthesised using the Neurosynth platform (Yarkoni *et al.*, 2011), we additionally draw facial landmarks for the purpose of visualisation. x,y,z values represent MNI coordinates and D: Dorsal, V: Ventral, A: Anterior, P: Posterior, L: Left, R: Right.

The second region is the *right PreFrontal Cortex (PFC)* which is also one of the most frequently reported “stress” activated regions (Wang *et al.*, 2005; Causse *et al.*, 2017; Rosenbaum *et al.*, 2018). The *right PFC*, in particular the *right dorsolateral PFC*, has been known to be associated with emotions (Knoch *et al.*, 2006). Compared with the left inferior frontal gyrus (Figure 7.1), the right PFC seems more suitable to be monitored from the top view of thermal imaging as it is located near the top of the scalp (the x -axis images in Figure 7.2). In Figure 7.2, we additionally segment the z -axis brain images (“ $z=+10$ ”, “ $z=+14$ ”) into three (vertical: anterior, middle, posterior) x two (horizontal: left, right) areas in order to facilitate the localisation of each region on a scalp thermal image (it shows the region is located in the right anterior area).

Other “stress” activated regions we found from the literature are the *occipital cortex* (involved in the visual perception) and the *right inferior parietal lobules* (involved in the perception of emotion and mathematical and linguistic operations) in the occipital/parietal lobes (Taylor *et al.*, 1997; Pruessner *et al.*, 2008). Figure 7.3 shows both regions (which are closely situated to each other) are located near the scalp (from x -axis images with labels A and P). From

z-axis images with segmentation, we can see the area is located in the right posterior area. Beyond the three regions which have been frequently reported, the *left PFC* (Figure 7.4a) and the *left premotor area* (Figure 7.4b) were also activated in Pruessner *et al.* (2008)’s study. Interestingly, stress appears to induce deactivation of some regions – the *superior temporal gyrus* (Figure 7.5) was deactivated in Taylor *et al.* (1997) and Pruessner *et al.* (2008).

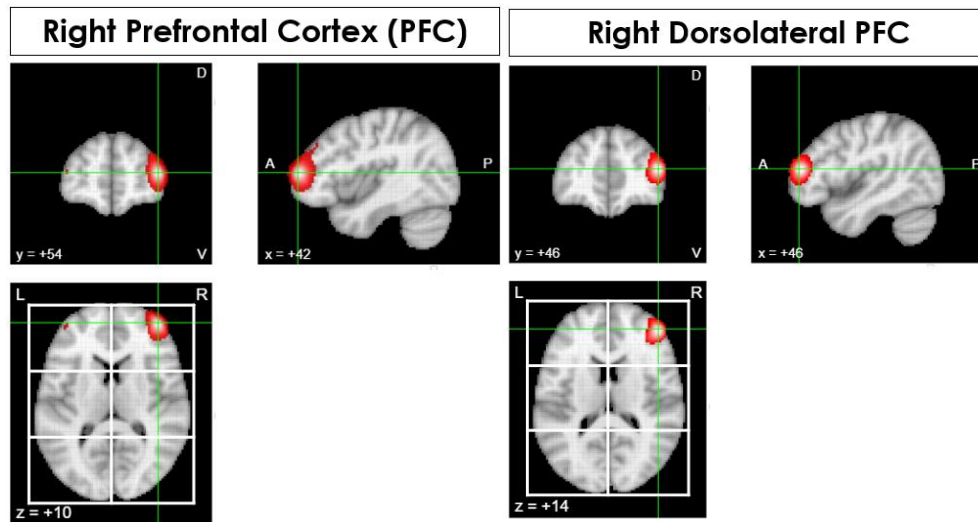


Figure 7.2. Location of the *right prefrontal cortex* and the *right dorsolateral prefrontal cortex*, the most commonly activated region under stress conditions (Wang et al., 2005; Causse et al., 2017; Rosenbaum et al., 2018): we segment the brain into 3 (anterior, middle, posterior) x 2 (left, right) areas on the *z*-axis images. *x,y,z* values represent MNI coordinates and D: Dorsal, V: Ventral, A: Anterior, P: Posterior, L: Left, R: Right.

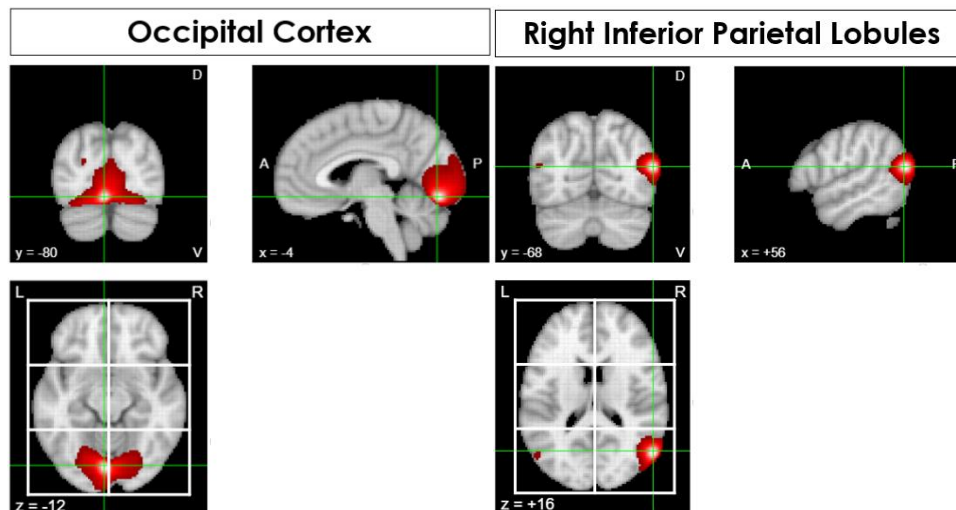


Figure 7.3. Location of the *parietal/occipital lobes* (the *occipital cortex* and *right inferior parietal lobules*), the frequently reported “stress” activation region (Taylor et al., 1997; Pruessner et al., 2008): we segment the brain into 3 (anterior, middle, posterior) x 2 (left, right) areas on the *z*-axis images. *x,y,z* values represent MNI coordinates and D: Dorsal, V: Ventral, A: Anterior, P: Posterior, L: Left, R: Right.

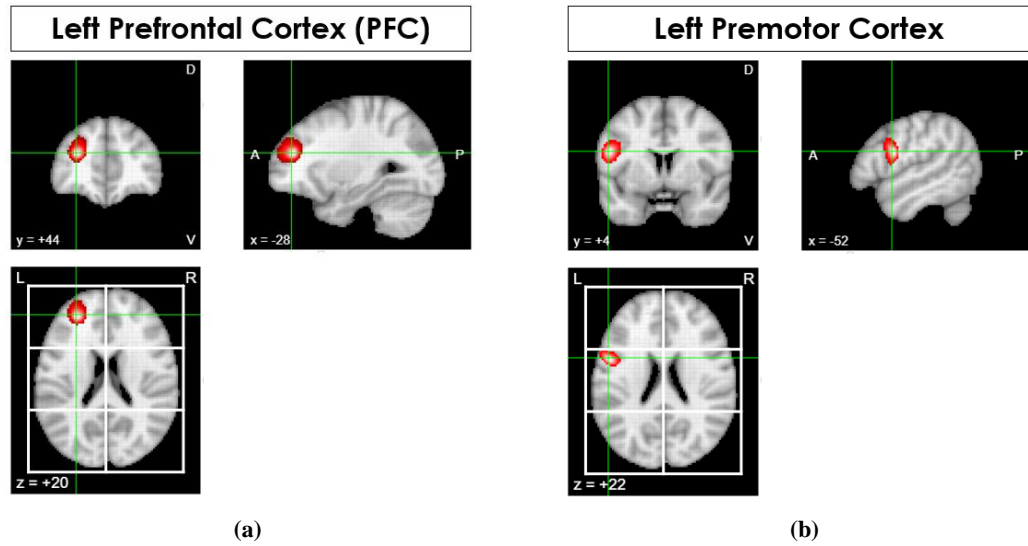


Figure 7.4. Localisation of the documented "stress" activation region: (a) the *left PFC* (Pruessner et al., 2008), (b) the *left premotor area* (Pruessner et al., 2008). x, y, z values represent MNI coordinates and D: Dorsal, V: Ventral, A: Anterior, P: Posterior, L: Left, R: Right.

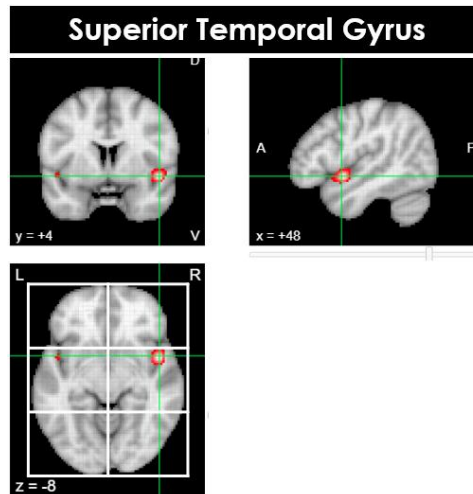


Figure 7.5. Localisation of the reported "stress" deactivation region: the superior temporal gyrus (Taylor et al., 1997; Pruessner et al., 2008). x, y, z values represent MNI coordinates and D: Dorsal, V: Ventral, A: Anterior, P: Posterior, L: Left, R: Right.

Thermal signal representation approaches

The stress-activated brain areas which have emerged from the literature have led to our research question: *can thermal responses of such brain regions to mental stress be observed by using low-cost thermal cameras? If so, how can we build methods and metrics to capture cortical thermal signatures?* To answer this, we review existing approaches (focusing on facial thermal

signatures) to represent thermal signals on a ROI, so as to build metrics to quantify novel cortical thermal signatures in our two studies. Schmidt documented earlier *“The average ... can be calculated with minimal cost. Calculating the average is meaningful for data from nearly any sensor”* (Schmidt, 2002). His comment reflects the fact that the average has been used as a major spatial representation method in studies using thermal imaging in order to capture spatial thermal trends of physiological information (Murthy & Pavlidis, 2006; Fei & Pavlidis, 2010; Abbas *et al.*, 2011; Lewis *et al.*, 2011; Gane *et al.*, 2011; Pavlidis *et al.*, 2012; Pereira *et al.*, 2015; Abdelrahman *et al.*, 2017).

Based on averaging, most authors have simply computed a difference between two selected mean temperatures as a metric (e.g. from baseline and from the final minute of a task in Abdelrahman *et al.*, 2017, or from the start and the end of a task session in Or & Duffy, 2007). A few authors have also applied basic statistical functions to a mean temperature time series extracted from sequential thermal image frames (e.g. the range, mean, variance, skewness of the mean temperature data in Gane *et al.* 2011). However, as we discussed in Chapters 5 and 6, the average-based spatial representation method and simple metrics built upon this may not be sufficient to capture complex aspects of thermal signatures, particularly, from the scalp. Thus, in this chapter we investigate potential metrics for capturing complicated cortical thermal signatures.

7.2. Study I: Initial Exploration in Unconstrained Setting

This section aims to undertake an initial exploration of cortical thermal signatures by conducting a study using the mental arithmetic task-based protocol used in Chapter 6. In particular, we hypothesise that *if a person perceives mentally stressed, his/her scalp area above (around) the documented stress activated brain regions will produce thermal directional changes and thermal imaging will help capture them*. As in Chapter 6, this study was conducted in an indoor unconstrained setting (we conduct our second study in a highly constrained setting in order to consider limitations found from this study).

We propose a thermal recording set-up, introduce an experimental protocol for this study and then discuss the collected dataset, evaluation method and results.

7.2.1. Recording Set-up

In order to observe stress-induced thermal signatures from the scalp, we use two low-cost, mobile thermal cameras (FLIR One Pro in Table 2.1) as the main sensing channels of this work. Figure 7.6 shows the proposed thermal recording set-up to monitor cortical thermal signatures. The first low-cost thermal camera is placed on a desk in front of a person to monitor the scalp from the frontal view (distance: 40-70cm, to enable participants' natural movement and postural changes). This is to primarily observe the scalp near the left inferior frontal gyrus (Figure 7.1) whose CBF is activated under stressful conditions, as discussed earlier. In this case, the camera can be directly attached to a smartphone (which controls the camera and stores data). The second thermal camera is installed above a person's head using a tripod and an additional mount (wire-based) to observe the scalp from the top view (distance: 40-60 cm). This is to investigate the other documented "stress activation" regions (i.e. left, right PFC, occipital/parietal lobes). Here, we take advantage of the lightweight portability of mobile thermal imaging in the installation. To reduce the weight further, an extension cable can be used to connect the camera to the smartphone (this setup helps a mobile thermal imaging device to be attached anywhere).

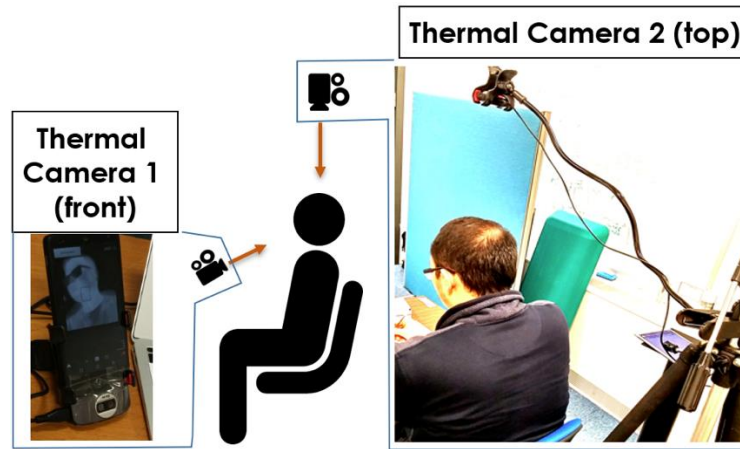


Figure 7.6. The proposed measurement architecture to monitor cortical thermal signatures. A low-cost thermal camera (directly attached to a smartphone) placed on a desk is to monitor the scalp from the front view. Another low-cost thermal camera (connected to a smartphone via a cable) is installed above the head of a person to monitor it from the top view.

7.2.2. Experimental Protocol and Dataset (*ThermalBrain I*)

This section describes the experimental protocol and the collected dataset. Amongst stress induction protocols used in neuroscientific experiments, the most frequently used task is mental arithmetic (Soufer *et al.*, 1998; Dedovic *et al.*, 2005, 2009; Wang *et al.*, 2005; Pruessner *et al.*,

2008; Al-Shargie *et al.*, 2016; Rosenbaum *et al.*, 2018). Hence, we used the mathematical serial subtraction protocol introduced in Chapter 6 where stress-induced thermal variability on the nasal area was explored. Although the protocol followed Soufer *et al.* (1998)'s introduction of time pressure in their PET study (i.e. 7.5 second for each mathematical question), mental arithmetic itself lacks other types of stressors (e.g. psychosocial stressors as addressed in Dedovic *et al.* 2009) so our protocol also included social evaluative threats and negative sound feedback as discussed in Section 3.3. Figure 7.7 describes the flow chart of the conducted study where the counting was used as the controlled session and the subtraction was used as the experimental session. In this data collection study, we collected thermal images from both cameras during the control and experimental sessions as the relaxation and control session induced significantly low stress levels, similarly to the experiment in Chapter 6.

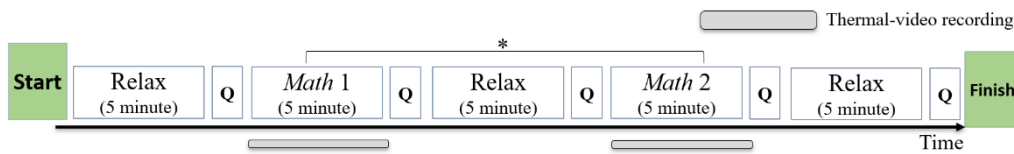


Figure 7.7. Flow chart of the study protocol: stress-induction task using mathematical serial subtraction (*Counterbalanced in Latin squared design).

We invited 10 participants (2 females with very long hair: 30cm~, 1 bald male and 7 males with short hair: 5cm ~ 10cm) who had not experienced the tasks before with a mean age of 31.2 (SD=11.67). Before conducting the experiment, participants were given the information sheet and consent form, as well as an additional demographics form used to check for any history of mental disease (see Appendix H) which could influence neural responses to mental stressors (Soufer *et al.*, 1998). Based on participants' self-report, no one had any history of mental diseases such as neurological disorder, substance abuse, head injury and psychiatric disorder, as well as severe physical illnesses. As shown in Section 6.3, each participant was instructed to rate subjective stress scores on a 10-cm Visual Analogue Scale (VAS). Any constraints such as limiting head movements were not imposed on participants in order to identify challenges in this unconstrained setup. The total experimental time for each participant was about 40-50 minutes. This produced a dataset of thermal images – we named this *ThermalBrain 1* dataset (thermal sequences from 2 sessions x 5 minutes x 10 participants).

7.2.3. Evaluation Method

Firstly, we tested the significance of subjective stress self-reports (VAS-based scores) to evaluate the study design (as conducted in Chapter 6). For this, we checked data normality and used a one-way repeated measures analysis of variance (ANOVA) with an independent variable: a session type with three conditions (resting, control and experimental).

The main aim of this study was to make an initial investigation of underexplored cortical thermal signatures near the documented stress-related local regions using the widely used metric to quantify the thermal signature. The metric is the difference between average temperatures at two temporal points. Motivated by our findings of the limitation of spatially averaging temperatures in Chapter 5, we further used two basic statistical functions (maximum, skewness), and one information-theoretic measure (entropy) for spatial representation. Based on the chosen four representation methods, we computed the difference between each spatially represented data from the start and end of a session. The four metrics used in this study are summarised in Table 7.1.

Table 7.1. Metrics for thermal directional changes based on four basic statistical measures (Average, Max, Skewness, Entropy) to represent thermal signatures from a ROI.

Metrics using Different types of Thermal Signature Representation		
Metric N1	Difference between [the data in the right column] from the start and the end of a session	Mean of 2D temperature elements on a ROI (*averaging: the most widely used representation)
Metric N2		Maximum of 2D temperature elements on a ROI
Metric N3		Skew of 2D temperature elements on a ROI
Metric I1		Entropy of 2D temperature elements on a ROI

We applied the four metrics to five scalp sub-regions near the *occipital lobe*, *left, right PFC*, *left premotor*, and *left inferior frontal gyrus* which all are CBF-activated regions in response to mental stressors. The computation of each value was manually done without the use of the ROI tracking algorithm used in earlier chapters. Paired-samples t-tests with a session type as IV (two conditions: control and experimental) were used to identify significant effects of mental stress (represented by the self-reported scores) on the thermal metrics from each region.

7.2.4. Results

Figure 7.8 shows examples of thermal images taken from both thermal cameras according to the set-up in Figure 7.6. Given a lack of registration methods for this, the bounding boxes were manually placed on the five ROIs (i.e. near the *left inferior frontal gyrus* (Figure 7.8a and 7.1) and near the left, right PFC, left premotor, and occipital lobe (Figure 7.8b and Figure 7.2-7.4), to extract the four metrics for thermal directional changes. We calculated each differential value from the first and last frame of thermal images. Results obtained 40 metrics values (5 ROIs x 2 main tasks x 4 metrics) from each participant.

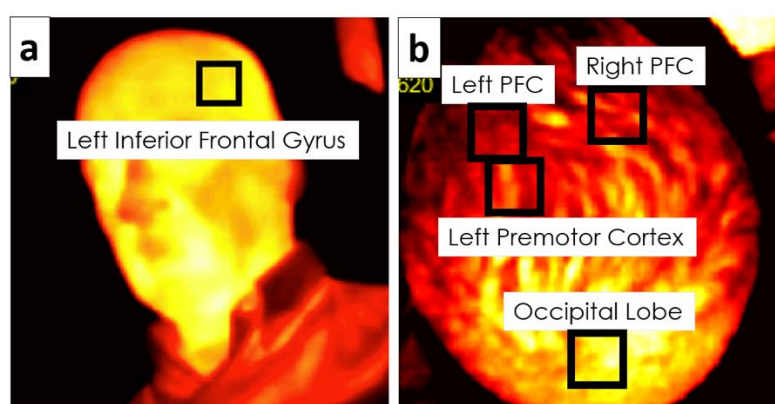


Figure 7.8. Examples of thermal images taken from (a) the frontal view thermal camera and (b) the top view thermal camera. The bounding boxes are placed on main ROIs: (a) near the left inferior frontal gyrus and (b) near the left, right PFC, left premotor, occipital lobe.

Figure 7.9 and 7.10 show examples of thermal images of the scalp collected from the experimental condition (Figure 7.9a, 7.10a) and the controlled condition (Figure 7.9b, 7.10b) in a chronological order. In the case of people with no hair (Figure 7.9), regions near the left prefrontal cortex (Figure 7.9a-1, top left circle) and occipital/parietal lobes (Figure 7.9a-8, bottom middle circle) showed a considerable increase in temperature under the stressful condition. Although the patterns were similarly shown on the thermal images of a participant with short hair, shown in Figure 7.10, we found that *the hairs interfered with monitoring temperatures of the scalp*. In particular, amongst the five ROIs, the scalp region above the *left inferior frontal gyrus* was hidden by very long hair of two female participants, which meant the data in the analysis had to be discarded.

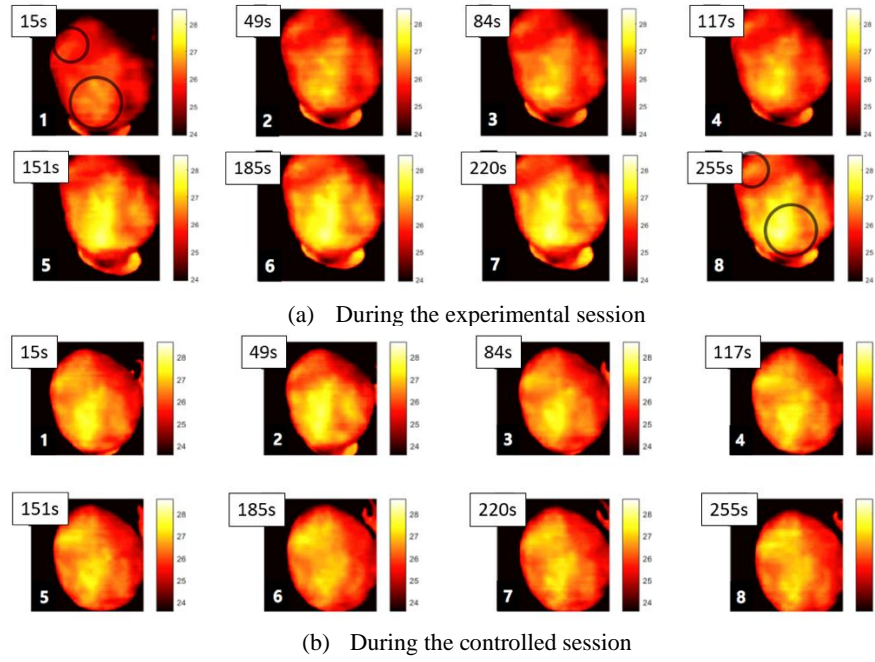


Figure 7.9. The scalp thermal imaging samples of a bald participant taken during (a) the experimental condition (i.e. the arithmetic subtraction and social evaluative pressures discussed in Section 3.3) and (b) the controlled condition (i.e. the arithmetic counting). Two circles were drawn to highlight temperature changes from the two regions (around left PFC and parietal/occipital lobes) in (a).

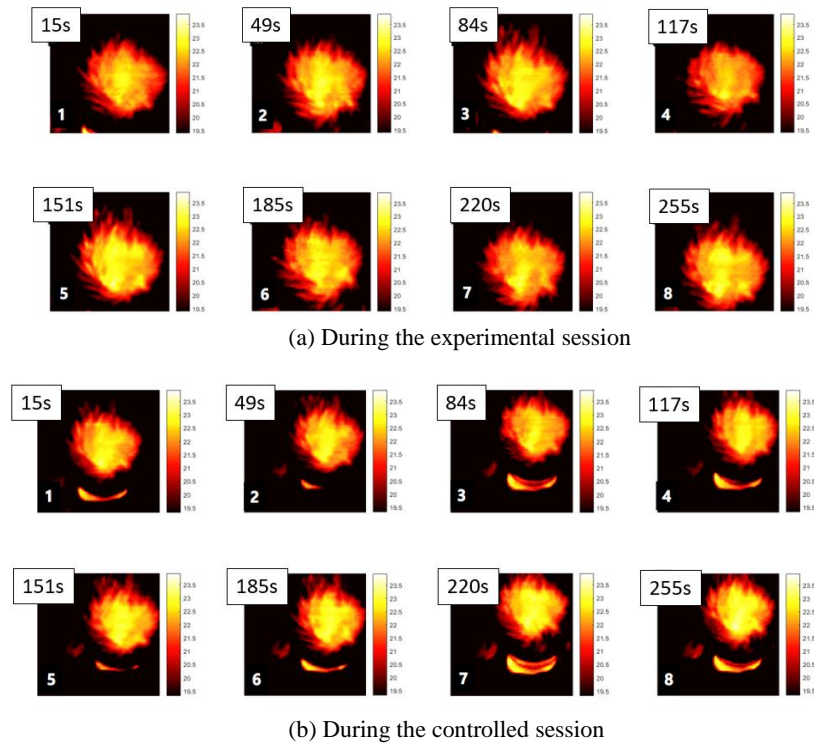


Figure 7.10. The scalp thermal imaging samples of a participant with a short hair taken during (a) the experimental condition (i.e. the arithmetic subtraction with social evaluative threat) and (b) the controlled condition (i.e. the arithmetic counting).

As in brain imaging studies with fMRI, PET (Pruessner *et al.*, 2008), where neural responses of participants during stress-induction sessions are directly compared with responses during controlled sessions, we analysed subjective ratings collected from ten participants to see if the stress induction protocol was properly designed to induce considerable differences of mental stress levels. The Shapiro-Wilk test confirmed that perceived self-reported scores for each condition were approximately normally distributed (resting: $p=0.379$, control: $p=0.464$, experimental: $p=0.233$). Using one-way repeated ANOVAs with a session type as IV, we found significant difference across each session (i.e. the controlled session using counting, experimental session using subtraction and the resting period) ($F(2,18)=16.458$, $p<0.001$, $\eta_p^2=0.646$). From the Bonferroni adjusted post-hoc analysis, the controlled condition is not significantly different to the baseline (i.e. resting period) ($p=0.743$), while the experimental condition is significantly different from both the baseline and control condition ($p=0.006$ with the controlled condition; $p<0.001$ with the resting condition).

For the significance test on the metrics, thermal recordings from one participant were excluded due to his *considerable head movements making it difficult to even manually localise and register the same area* on each thermal image, in particular from the top view, as the main ROIs on the participant's head were not visible (due to the angle). This indicates that, for a better *registration of each ROI* in the follow-up study, we may need to limit a participant's head movement or build a method that computationally aligns the orientation and position of the scalp on a thermal image with others. Additionally, the recordings from two participants were also excluded as we found *they gave up the subtraction task due to the extreme difficulty they experienced* from the brief interview done after the end of all sessions, indicating a need to *adjust the difficulty levels* according to participants. This exclusion also led to data normality.

Paired-samples t-test (IV: a session type with two conditions) were used to analyse effects of mental stress on each thermal metric value (see metrics in Table 7.1). Figure 7.11 visualised a summary of the significance test results. We found significant differences between the stress-induction experimental session (i.e. arithmetic subtraction) and controlled one (i.e. counting-down) from the region over the *left inferior frontal gyrus* (N1: the difference of the mean data, $t(4)=4.964$, $p<0.01$; N2: the difference of the maximum data, $t(4)=2.849$, $p<0.05$) in Figure 7.11a and over the *occipital lobe* (N1: $t(6)=4.34$, $p<0.05$; N2: $t(6)=2.85$, $p<0.05$) in Figure 7.11b. As illustrated in Figure 7.12 (Left, Middle), both regions produced heat during the stress-induction task, while they tended to lose heat in the controlled condition. Furthermore, we found an approaching significance result for the skewness-related metric (N3) applied to the region over

the *left premotor area* ($t(6)=2.45$, $p=0.05$) with an increase in skews under the experimental condition, as shown in Figure 7.11c.

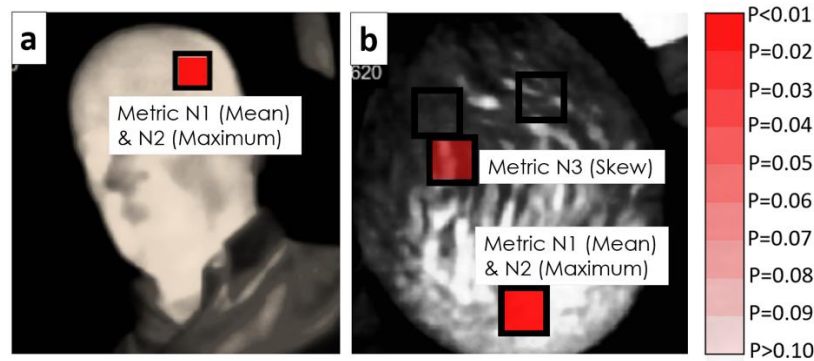


Figure 7.11. Significance test results visualised on each thermal image from (a) the frontal view thermal camera and (b) the top view thermal camera. The bounding boxes are placed on main ROIs: (a) near the left inferior frontal gyrus and (b) near the left, right PFC, left premotor, occipital lobe.

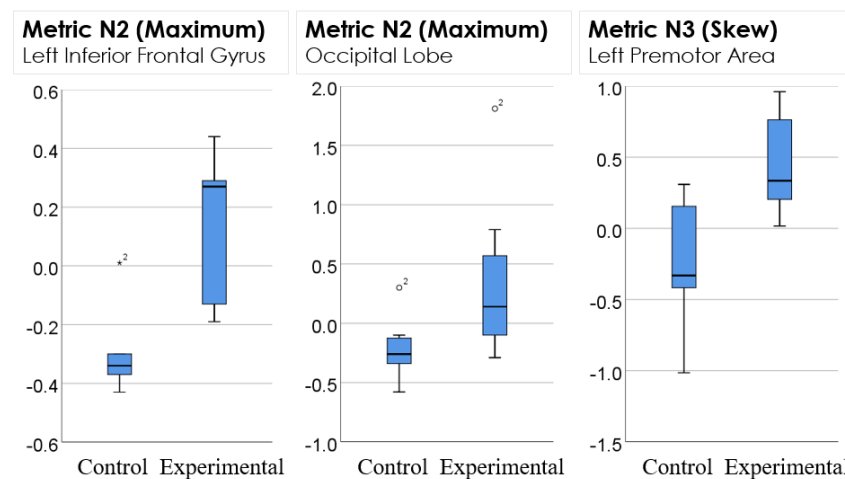


Figure 7.12. Box plots of 95% confidence intervals in selected (examples) metrics values across each session (control vs. experimental). Left: Metric N2 (the difference of the maximum data) from the left inferior frontal gyrus, Middle: Metric N2 from the occipital gyrus, Right: Metric N3 (the difference of the skewness data) from the left premotor area.

7.2.5. Reflection on Study I: Identified Challenges

The results are encouraging as we have found initial evidence of measurable thermal propagation from CBF-activation regions using low-cost thermal cameras. However, we have also identified challenges in this study protocol: i) long hair hampering the scalp temperature, ii) head movement

inducing registration difficulty, iii) extreme difficulty of mental arithmetic leading to a participant giving up. In addition, we have found other types of spatial representation methods beyond averaging (e.g. maximum, skewness-based methods in Table 7.1) that could be used for building metrics for quantifying cortical thermal signatures. Given these initial results, in the follow-up study in Section 7.3, we design and conduct a highly controlled study in terms of the three challenges and investigate metrics for improving the capture of cortical thermal signatures.

7.3. Study II: London Arithmetic Stress Task (LAST) in Constrained Setting

Based upon the identified challenges and findings from the first study in Section 7.2, this section aims to undertake an in-depth investigation on cortical thermal signatures by introducing a highly constrained setting, limiting hair length and head movement. In addition, we introduce a more refined study protocol built on the mental arithmetic task by following an exemplary experimental protocol used in neuroscientific studies on mental stress. As in Study I, the key hypothesis in this study is: *if a person perceives mentally stressed, his/her scalp area above (around) the documented stress activated brain regions will produce thermal directional changes and thermal imaging will help capture them*. Subsections below present the methods, materials and collected dataset. This section ends with reporting the results.

7.3.1. Recording Set-up

Figure 7.13 shows the proposed recording set-up to monitor cortical thermal signatures. This setup is primarily designed to limit head motions (one of the identified challenges) of each participant, and to support the registration of each local region. As shown in Figure 7.13a and 7.13b, a chin rest is installed on a desk where a laptop is placed. Instead of using one top-view thermal camera in Figure 7.6, here we introduce two top-side (top-left, top-right) view thermal cameras (the lost cost, mobile ones) to monitor each part of the scalp over left and right hemispheres of the brain more closely. Because we use the chin rest (with two poles) which interferes with the frontal view, we only focus on the two top-side view in this experiment. As seen in Section 7.2.1, a wire-based mount and an extension cable (connecting a smartphone and a camera) are used to support the installation of each camera above the scalp.

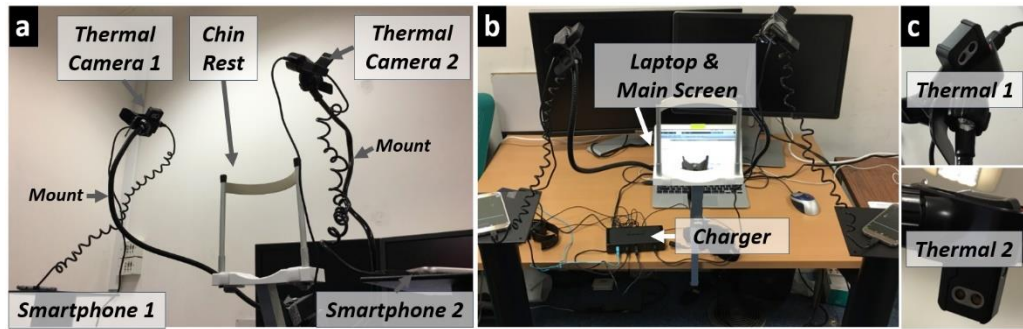


Figure 7.13. The proposed measurement architecture to monitor cortical thermal signatures. To limit each participant's head motion, (a) a chin rest is installed on (b) a desk where a laptop is placed. (c) Two low-cost thermal cameras are installed above one's head to monitor (than the setup in Figure 7.6) each part of the scalp over left and right hemispheres of the brain more closely.

Figure 7.14 illustrates how each thermal camera in Figure 7.13 captures the scalp temperature. With this setup, although it is difficult to capture temperatures over the left inferior frontal gyrus due to the chin rest hiding the area, we can capture better close-up images of each part of the scalp over the left and right hemispheres in comparison with the setup used in the first study (Section 7.2). Following Lombard's earlier approach (1878), the red circle segmented into three subregions (anterior, middle, posterior) is used to register local regions on each hemisphere. This also helps to directly compare the three (vertical: anterior, middle, posterior) by two (horizontal: left, right) segmented areas on brain images in Figure 7.1-7.5 where the documented CBF activation/deactivation regions are highlighted. More details of the data extraction are described in Section 7.3.2.

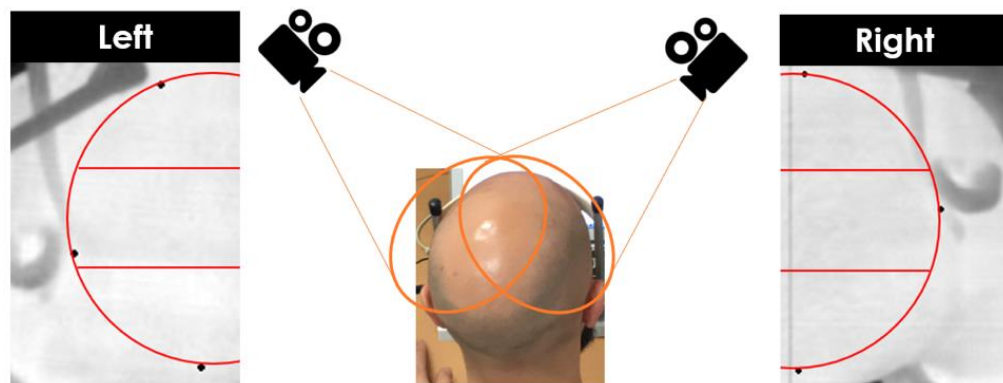


Figure 7.14. Examples of thermal images capturing the scalp over the left and right hemispheres. The red circle with two lines inside the thermal images segments the scalp region into 6 subregions following Lombard's earlier approach (1878). More details are described in Section 7.3.2.

7.3.2. Proposed Computational Methods

Using the measurement architecture proposed in Section 7.3.1, we propose a new method to extract time-varying cortical signatures from scalp thermal images and new metrics to quantify cortical thermal activation and its variability.

7.3.2.1. Extraction of Time-varying Cortical Signatures

Figure 7.15 illustrates a series of computations we propose for extracting temporal (time series) cortical thermal signatures. It consists of four steps: i) motion correction, ii) registration, iii) segmentation, and iv) cortical signature representation. In Step I, despite the use of a chin rest to limit a person's head motion, we use the ROI tracking algorithm (the Optimal Quantisation-based Thermal Gradient Flow in Chapter 4 and 5) in order to correct even a very subtle motion artefact which could lead to noisy signals. For this, we chose the entire head as a ROI to track. Step II is to register each hemisphere (right and left) on thermal images. This is done by choosing three points around each hemisphere to draw a circle using the Hough transform (Duda & Hart, 1972). In Step III, following Lombard (1878), we segment the circular area into three local areas (A: anterior, M: middle, P: posterior). The anterior area roughly includes the PFC, the middle and the posterior areas include the premotor/primary motor areas and parietal/occipital lobes, respectively. The final step (Step IV) is to extract cortical thermal signatures from each segmented region using basic statistical functions (maximum, average, skewness). An example pair of spatial skew signals from the three segmented areas is shown in Figure 7.15 (Step IV, right): we collected skew signals from a participant's right hemisphere during a stress-induction section. The signal from the right PFC area (A in the figure) was highly fluctuated, possibly reflecting the areas response to mental stress.

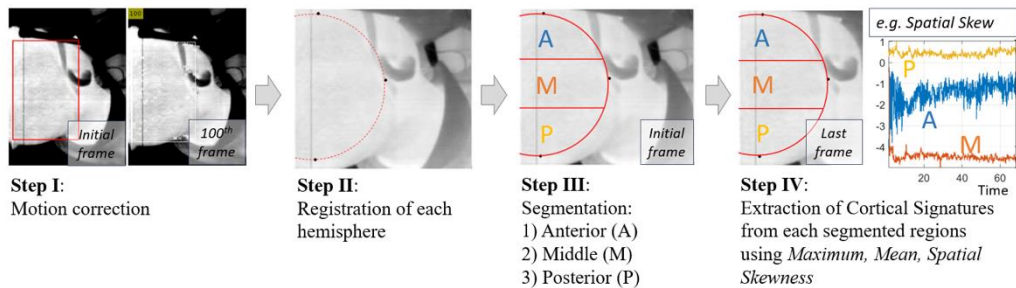


Figure 7.15. The proposed method for extracting time-varying cortical signatures from the scalp: Step I) motion correction by selecting the scalp as a main ROI and using a ROI tracking algorithm, Step II) registration of each hemisphere by choosing three points to draw a circle (using the Hough transform), Step III) segmentation into three areas (A: anterior, M: middle, P: posterior) following Lombard (1878), Step IV) extracted cortical thermal signatures from each segmented region using basic statistical functions (maximum, average, skewness).

7.3.2.2. Metrics for Quantification of Cortical Thermal Signatures

As explored in Section 7.2, not only the average, however also the skewness and maximum measurements can help to represent cortical signatures from a ROI on a thermal image. Extending the metrics (N1-N3) in the initial investigation (Table 7.1), we propose a set of basic metrics to quantify cortical thermal activation and deactivation. Similarly to N2, we compute the difference between *the maximum data from the end of a session (control or experimental)* (for example, the value from the posterior area (*P*) at 170s in Figure 7.16 bottom) and *the mean of the maximum time series from a baseline period* (for example, the average of timeseries from the *P* area in the baseline period in Figure 7.16 top). The same can be applied to the mean and skew, building metrics N4-6 as in Table 7.1.

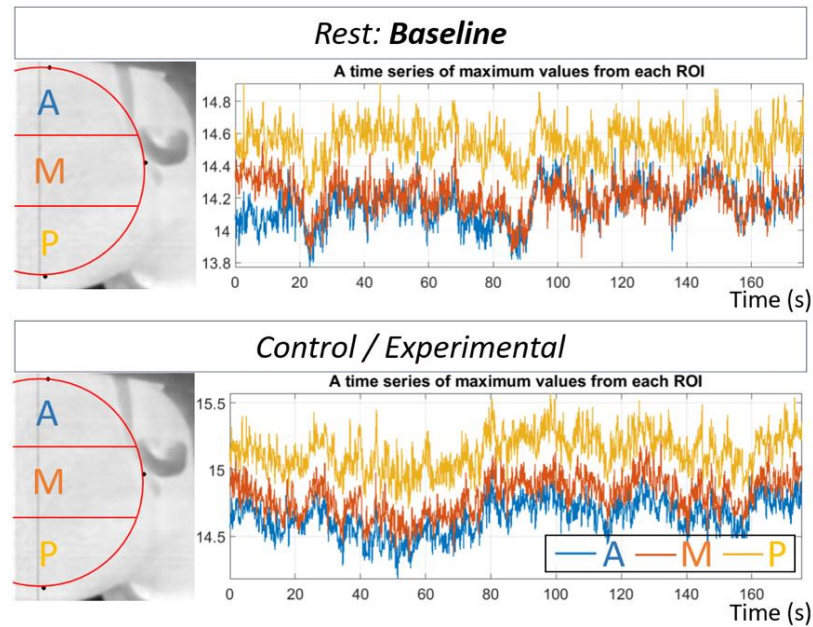


Figure 7.16. Examples of a maximum-based time series from each ROI. Top: from a baseline, Bottom: from a controlled or experimental session.

As used in Engert *et al.* (2014), we compute the slope (using a linear polynomial fitting) from the mean, maximum, skew timeseries, building metric N7-9. As in Chapter 6 we have demonstrated the key ability of feature-scaling in handling physiological interpersonal differences, we apply feature scaling (z-score based) to the nine metrics N1-9, resulting in another set of nine metrics N10-18. Normalisation using baseline data (Vizer *et al.*, 2009) was also considered to build the last nine metrics N19-27. The discussed basic metrics are summarised in Table 7.2.

Table 7.2. A set of basic metrics to quantify cortical thermal signatures.

Basic metrics for cortical thermal signatures		
Metric N1	Difference between [the data in the right column] from the start and the end of a session	Mean (of 2D thermal data on a ROI)
Metric N2		Maximum
Metric N3		Spatial Skew
Metric N4	Difference between [the data in the right] from baseline (mean of the data) and the end of a session	Mean
Metric N5		Maximum
Metric N6		Spatial Skew
Metric N7	Slope (using a linear polynomial fitting) of [the time series data in the right]	Mean
Metric N8		Maximum
Metric N9		Spatial Skew
Metric N10	Difference between [the data in the right column] from the start and the end of a session	Feature scaled Mean (across a session)
Metric N11		Feature scaled Maximum
Metric N12		Feature scaled Spatial Skew
Metric N13	Difference between [the data in the right] from baseline (mean of the data) and the end of a session	Feature scaled Mean
Metric N14		Feature scaled Maximum
Metric N15		Feature scaled Spatial Skew
Metric N16	Slope (using a linear polynomial fitting) of [the time series data in the right]	Feature scaled Mean
Metric N17		Feature scaled Maximum
Metric N18		Feature scaled Spatial Skew
Metric N19	Difference between [the data in the right column] from the start and the end of a session	Normalised Mean (using baseline data)
Metric N20		Normalised Maximum
Metric N21		Normalised Spatial Skew
Metric N22	Difference between [the data in the right] from baseline (mean of the data) and the end of a session	Normalised Mean (using baseline data)
Metric N23		Normalised Maximum
Metric N24		Normalised Spatial Skew
Metric N25	Slope (using a linear polynomial fitting) of [the time series data in the right]	Normalised Mean (using baseline data)
Metric N26		Normalised Maximum
Metric N27		Normalised Spatial Skew

However, it is expected that the proposed basic metrics N1-N27 in Table 7.2 would not be very sensitive to variability in thermal signatures as discussed in Chapter 6. Hence, we develop a set of metrics to quantify cortical thermal variability in the mean, max, skew time series. In particular, it has been shown that the skewness measurement can be sensitive at detecting a differentiated state of tissue composition on a local ROI in other domains, such as breast cancer detection in breast x-ray imaging (Byng *et al.*, 1996). Inspired by this, we assume that skewness can be a good measure of spatially changing patterns in CBF-induced thermal propagation on the scalp. Spatial skew $\gamma_1(k)$ at k^{th} frame can be measured by

$$\gamma_1(k) = E \left[\frac{T_{ij}(k) - \mu_k}{\sigma_k} \right]^3. \quad (7.1)$$

where $T_{ij}(k)$ is a temperature element at (i, j) in the x-y plane (a two-dimensional space) from the k^{th} frame.

To quantify variability of the spatial skewness timeseries, we propose a series of pre-processing steps and measures based upon it as summarised in Figure 7.17. The first step is to resample the signal to address the unsteady frame rate of the low-cost thermal camera, as addressed in Chapters 5 and 6. For the implementation, we used a linear interpolation to resample signals (16Hz). Here, with the assumption that the quality of such time-varying signals is affected by low- and high-frequency noise (e.g. motion artefacts) similar to fNIRS data (Rosenbaum *et al.*, 2018), we perform a bandpass filtering (0.01-0.1 Hz) as in (Rosenbaum *et al.*, 2018). Finally, using the filtered signal, we propose two measures to capture variability: the first is a square root of the number of zero crossing points (with respect to a border between negative-positive skews) and the second is a square root of the number of local maxima. For counting local maxima, we simply used a Matlab function *findpeaks* with a condition of 10% minimum peak prominence (empirically chosen). Additionally, we apply both measures to the mean, maximum time series, leading to four metrics (V3-V6). The proposed metrics are summarised in Table 7.3.

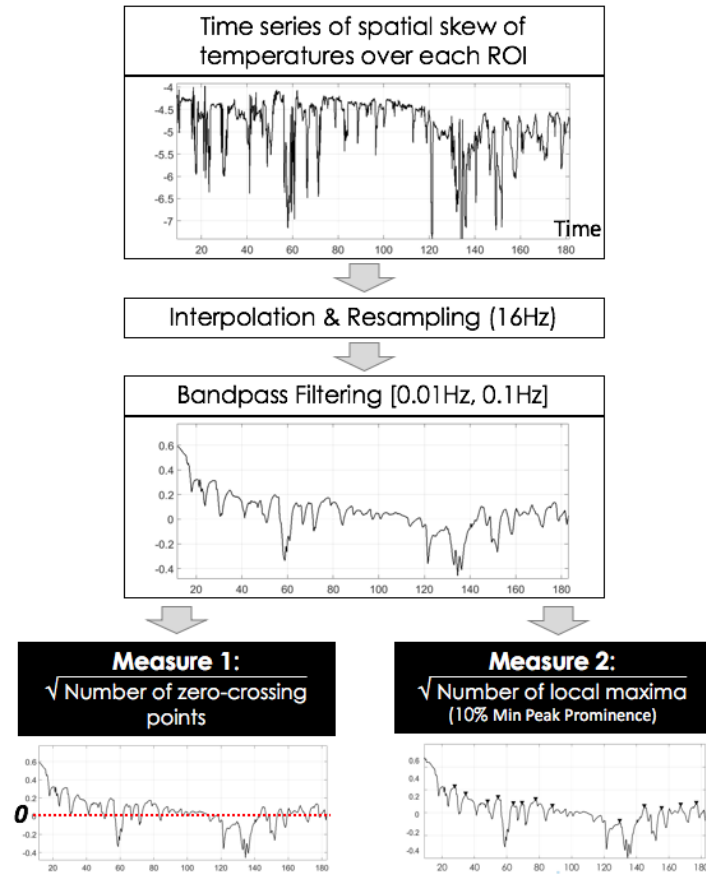


Figure 7.17. Pre-processing (interpolation & resampling and bandpass filtering) for building measures to quantify variability of the skewness timeseries from a ROI. The first measure is the number of zero-crossing points, and the second is the number of local maxima using a Matlab function *findpeaks* with a condition of 10% min peak prominence.

Table 7.3. A set of metrics to quantify variability of cortical thermal signatures.

Proposed metrics to quantify cortical signature variability		
<i>Metric</i>	<i>Measures</i>	<i>Targeted time series data</i>
Metric V1	Number of zero-crossing points	Spatial Skew
Metric V2	Number of local maxima	
Metric V3	Number of zero-crossing points	Mean
Metric V4	Number of local maxima	
Metric V5	Number of zero-crossing points	Maximum
Metric V6	Number of local maxima	

7.3.3. Experimental Protocol and Dataset (*ThermalBrain II*)

This section describes the experimental protocol and the collected dataset. With the proposed recording set-up (Figure 7.13), helping to minimise motion artefacts, the study aims to explore complicated aspects of cortical thermal signature (beyond thermal directionality) in highly controlled settings. We design this protocol built on the protocol used in Section 7.2 (using mental arithmetic, social evaluative, time pressures, and negative sound feedback). Given a lack of ground truth measurements (because existing neural activity measurement tools interfere with collecting scalp thermal images), we also follow the timeline and conditions used in the exemplary neuroscientific experimental protocol (mental arithmetic-based) of Pruessner *et al.* (2008). The protocol in Pruessner *et al.* (2008) uses a block design with two repeated runs. Each run consisted of three sessions: i) rest, ii) control, and iii) experimental sessions. Each session lasts for 3 minutes and the beginning condition is counterbalanced between participants with the consistent sequence of the three sessions (e.g. ... - rest - control - experimental - ...), as shown in Figure 7.18. The timeline of the built task procedure, named as London Arithmetic Stress Task (LAST), is summarised in Figure 7.18a. Again, following Pruessner *et al.* (2008), we use one LAST block as a training session before the main LAST (Figure 7.18b).

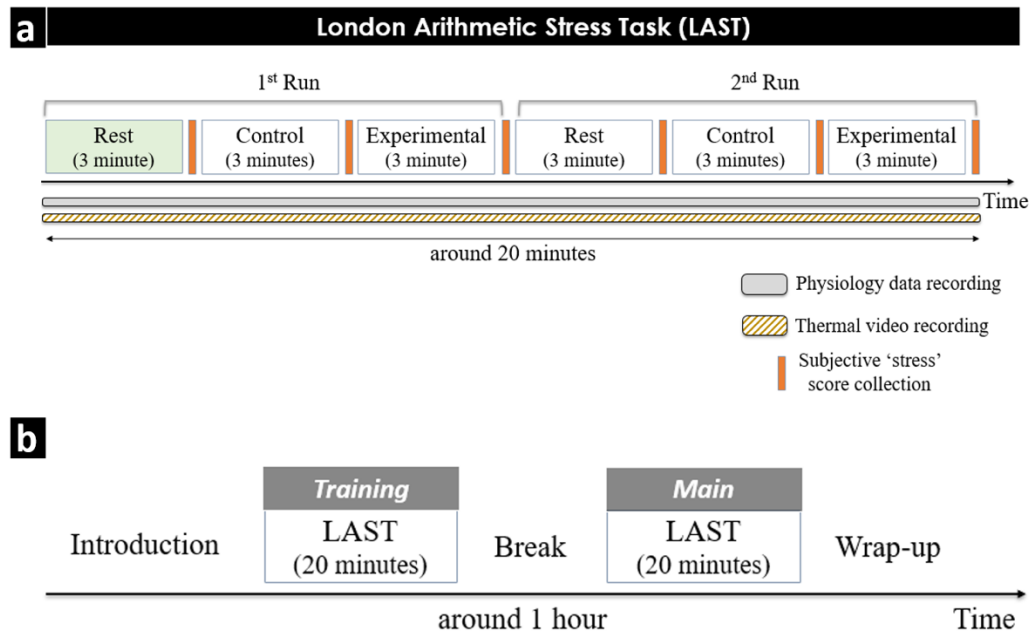


Figure 7.18. Timeline of the London Arithmetic Stress Task (LAST): (a) the LAST has two runs and each run consists of three conditions, rest (baseline), control, experimental over the course of the nine minutes as in Pruessner et al. (2008). While the beginning condition (rest, stress or control) is counterbalanced between participants, the sequence of the three conditions is kept constant between participants: control was following rest and preceding stress. (b) A training session is included in the entire LAST protocol as in Pruessner et al. (2008).

Considering the remaining challenges identified from Study I (i.e. *long hair hampering the scalp temperature, extreme difficulty of mental arithmetic leading to a participant giving up*), in this study we chose participants with a specific type of hair (bald or short) and modified our mental arithmetic question program so as to adapt its difficulty levels to each participant's performance. We programmed this to choose a subtractor from a set of prime numbers along with a participant's latest correct rates (e.g. the last three questions), i.e., if a user chooses wrong answers for three times in a row, the subtractor is set to a lower prime number, for example from 13 to 11. This can help participants remain fully engaged during the mental arithmetic task (Soufer *et al.*, 1998).

We invited 10 healthy male participants (aged 21-48 years, $M=34.4$, $SD=7.56$) of varying ethnicities and different skin tones (pale white to black) from the University College London and non-research community through the UCL psychology subject pool system. They all had very short hair (almost bald) which are summarised in Table 7.4 (visually inspected with a scale). Participants completed pre-screening through the system which was designed to exclude participants with any history of psychiatric disorders or medicine intakes which may influence

their cortical signatures. No one had any history of mental diseases like neurological disorder, substance abuse, head injury and psychiatric disorder as well as severe physical illnesses.

Table 7.4. Hair types of participants for Study II.

Participant	Hair Type
P1	<i>0.5-1cm (curly)</i>
P2	<i>0.5cm (curly) and almost bald</i>
P3	<i>0.5cm_almost bald</i>
P4	<i>2-3cm</i>
P5	<i>Almost Bald</i>
P6	<i>5-6cm</i>
P7	<i>Bald</i>
P8	<i>Bald</i>
P9	<i>Almost Bald</i>
P10	<i>Bald</i>

Every participant was asked to place their chin on the chin rest during each session as shown in Figure 7.19. Before the start of each run, we checked whether the chin rest was comfortable, otherwise we adjusted the height and the position of the equipment according to participant's needs. During the rest sessions, participants were asked to rest as they wanted to ensure they could relax. As in Study I, each participant was instructed to rate subjective stress scores on a 10-cm Visual Analogue Scale (VAS). As shown in Figure 7.19, with two top-view low-cost thermal cameras, an additional thermal camera was installed in front of a person; however, as explained in Section 7.3.1, the chin rest interfered with monitoring the area around the inferior frontal gyrus. We did not use the data from the frontal camera for the analysis. We also measured blood volume pulse, skin conductance, breathing signals using a finger PPG (ProComp Infiniti BVP/SA9308M, Thought Technology), two bracelet-based GSR/PPG sensors (worn on each left, right wrist, Empatica E4) and a breathing belt (ProComp Infiniti Resp/SA9311M, Thought Technology) in order to build a rich physiological dataset (however, the thesis only uses data from the two top-view thermal cameras for the analysis). The total experimental time for each participant was about 75-90 minutes. This produced a dataset of thermal images and physiological signals – we named this *ThermalBrain II* dataset.

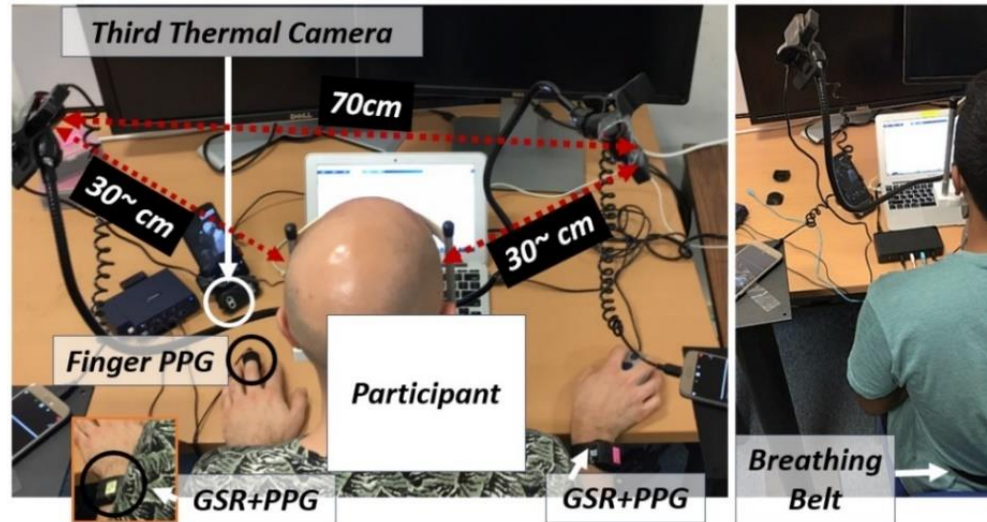


Figure 7.19. Experimental setup: with two top-view thermal cameras, an additional third camera was installed in front of a person. A finger PPG, two bracelet-based GSR/PPG sensors (worn on left, right wrists) and a breathing belt were used to build a rich physiological dataset.

7.3.4. Evaluation Method

In this study with a highly controlled setup, we would like to answer these questions: i) *Can we find significant cortical thermal changes from the documented “stress” regions when noise due to head motion and hair condition is controlled?*, ii) *Are the findings consistent over repeated tasks?*, iii) *if so, which metrics will play significant roles in capturing cortical signatures?*

Firstly, we conducted significance tests on subjective stress self-reports (VAS-based scores) to check if the protocol was properly designed in terms of whether the repetition influences perceived stress scores. We planned to use a two-way repeated measures analysis of variance (ANOVA) test with two Independent Variables (**IVs**): a task type (three conditions: baseline, control and experimental), repetition (two conditions: first and the second run). However, self-reported scores for two cases were skewed; thus, we instead conducted a Friedman’s ANOVA test with a task type as IV separately for the two runs. Then, we carried out the nonparametric Wilcoxon signed ranks test to compare each pair.

Secondly, we tested for effects of the two IVs (i.e. task type and repetition) on the proposed metrics applied to the six segmented areas as dependent variables. As metrics already use the baseline information, we removed the baseline from the condition for the first IV. Lastly, with the focus on the data from the first run, we tested significance of the task type on each metric to confirm any differences from the second run. As values from each metric were normally

distributed and met all assumptions (e.g. sphericity) for each case, we used the ANOVA test which is followed by pairwise post-hoc Bonferroni tests with adjustment (Girden, 1992) to compare the performance of the different algorithms.

7.3.5. Results

Analysis of perceived stress levels

From the stress-induction study, we collected 60 self-reported scores (10 participants x 3 sessions x 2 runs). The boxplot in Figure 7.20 shows the distribution of the self-reported scores over the resting, control and experimental session from each run. It is clear that the experimental condition (mental arithmetic, social evaluative threats, time pressure and negative sound feedback) generally induced much higher levels of stress than the resting condition while the control condition (counting down) did not (From the **first run**: Rest M=1.36, SD=1.21; Control M=1.67, SD=1.71; Experimental M=4.47, SD=2.10; From the **second run**: Rest M=1.16, SD=1.92; Control M=1.50, SD=1.32; Experimental M=4.53, SD=2.44).

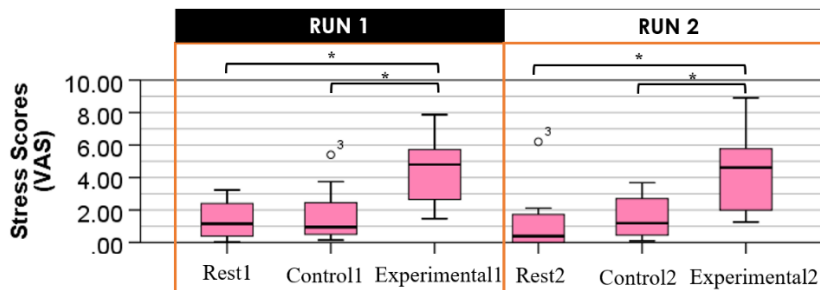


Figure 7.20. Boxplots (95% confidence interval) of self-reported stress scores from the 10 participants across each section (rest1, control1, experimental1 from the first run and rest2, control2, experimental2 from the second run).

As the data for Control1 and Rest2 were skewed ($p=0.026$, $p<0.001$ from the Shapiro-Wilk test), we conducted the Friedman test on the 10 participants' scores for each run separately. The results showed a significant effect of the task type on the perceived stress levels for both runs ($\chi^2(2)=14.6$, $p=0.001$), indicating that the task type is the main factor differentiating a participant's mental stress. Wilcoxon signed ranks test showed that the experimental condition induced significantly higher levels of mental stress than the two other sessions for each run (Run

1: Experimental – Baseline: $Z=-2.803$, $p=0.005$, Experimental – Control: $Z=-2.599$, $p=0.009$; Run 2: Experimental – Baseline: $Z=-2.803$, $p=0.005$, Experimental – Control: $Z=-2.701$, $p=0.007$), whilst the control condition did not show significant difference from the baseline condition (Run 1: $Z=-1.07$, $p=0.285$; Run 2: $Z=-0.968$, $p=0.333$), indicating that stressors used in the experimental sessions were the dominant factor inducing high-levels of mental stress. On the other hand, there were no significant effects of repetition (i.e. each run) on perceived mental stress levels for each task (Baseline: $Z=-1.478$, $p=0.139$; Control: $Z=-0.255$, $p=0.799$; Experimental: $Z=-0.051$, $p=0.959$), confirming that the study protocol was designed properly.

Analysis of proposed metrics in assessing mental stress

From 120 collected thermal images (10 participants x 3 sessions x 2 runs x 2 thermal cameras), we extracted 792 metrics values (33 metrics from Table 7.2 and Table 7.3 x 6 segmented local areas x 2 sessions x 2 runs: as data from baseline periods are used to compute these metrics values, here we have 2 sessions). Using the data, we tested the effect of a session type (control and experimental) on each metric value applied to each segmented local areas (left anterior, left middle, left posterior, right anterior, right middle, right posterior).

From a two-way repeated measures ANOVA analysis test with two IVs (task type and repetition), we found that the metric **V1** (based on zero-crossing points of the spatial skew time series) showed significant effects of task type on the right anterior ($F(1,9)=8.03$, $p=0.02$, $\eta_p^2=0.472$) and right middle areas ($F(1,9)=6.844$, $p=0.028$, $\eta_p^2=0.432$) as visualised in colour in Figure 7.21 (strong red: $p<0.01$, light red: $p>0.10$), whilst the same metric showed no significant effects of repetition and no significant interaction effects upon each area (the data for each combination of IVs were normally distributed with p-values ranged between 0.103 and 0.971 from the Shapiro-Wilk test). As shown in Figure 7.22, the metric **V1** value was significantly higher in the experimental condition than in the control condition for each run, which may indicate that stress-induced cortical activities lead to rapidly fluctuating thermal patterns (rather than just increasing or decreasing temperature).

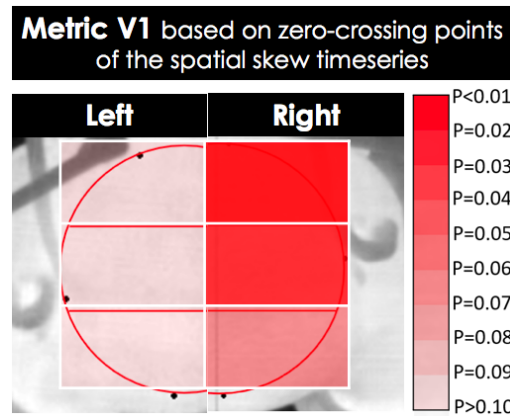


Figure 7.21. Results from significance tests of the task type over the two runs on the metric **V1** (based on zero-crossing points of the spatial skew time series) are visualised on each thermal image of the scalp over the left and right hemispheres. The right anterior and middle areas (above the right PFC and superior temporal gyrus) had significance.

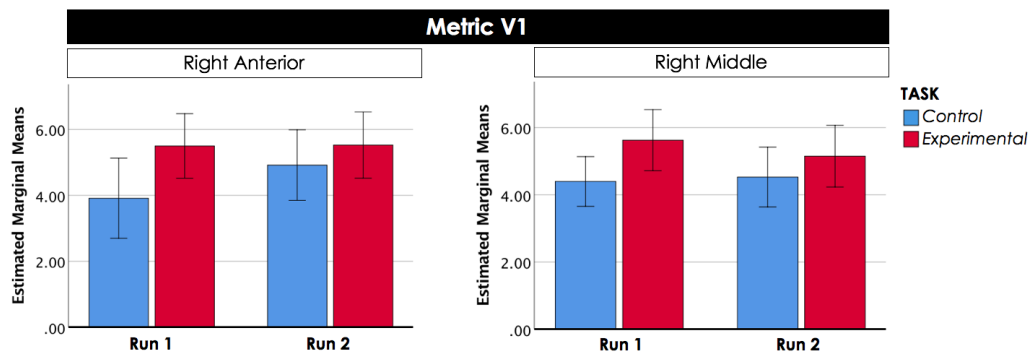


Figure 7.22. Bar plots of the metric **V1** values from the right anterior and middle areas across each run and each task condition.

Likewise, as visualised in Figure 7.23 and 7.24, we also found the metric **V2** (based on local maxima of the spatial skew time series) from the right anterior area had significant increases under the experimental condition ($F(1,9)=14.857$, $p=0.004$, $\eta_p^2=0.623$) and the metric on the right posterior area had an approaching significant effect of the task type ($F(1,9)=4.72$, $p=0.058$, $\eta_p^2=0.344$). There were no significant effects of repetition and no interaction effects of the IVs. The data for each case were normally distributed (p-values ranged between 0.118 and 0.929 from the Shapiro-Wilk test). On the other hand, all the other metrics (basic N1-N27 and variability V3-V6) applied to all six areas showed no significant effects for both task type and repetition.

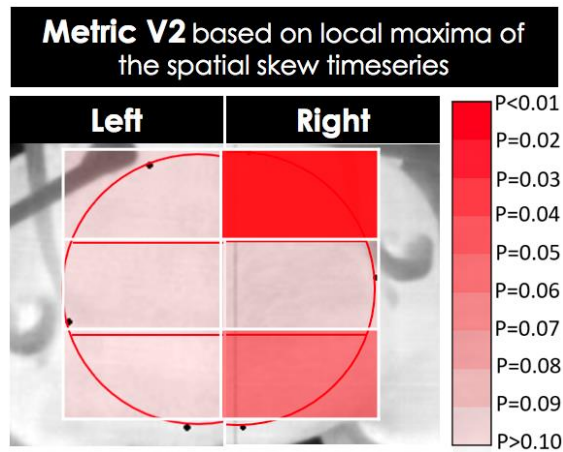


Figure 7.23. Results from significance tests of the task type over the two runs on the metric **V2** (based upon local maxima of the spatial skew time series) are visualised on each thermal image of the scalp over the left and right hemispheres. The right anterior (above the right PFC) had significance and the right posterior area (above the right parietal/occipital lobes) had an approaching significance.

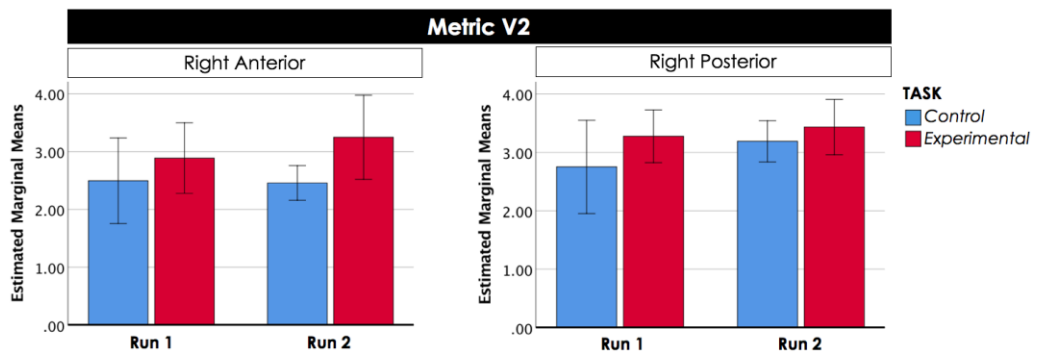


Figure 7.24. Bar plots of the metric **V2** values from the right anterior and posterior areas across each run and each task condition.

Lastly, in order to test effects of the repetition on the power of basic metrics which were designed to capture directional changes in each timeseries (which showed significant effects of the task type from some regions in Study I – e.g. the area above occipital lobe in Figure 7.11b), we additionally conducted a paired-samples t-test on the basic metrics N4,N5 (extended from N1,N2 - directional change in mean / maximum temperature timeseries- using baseline data) values from the first run and the second run separately. Here, the IV is the task type. Interestingly, the results showed both N1 and N4 (based on the average representation) from *the first run* showed a significant effect of the task type on the left anterior ($F(1,9)=5.172$, $p=0.049$), while those from *the second run* showed no significant effect of the task type on it ($F(1,9)=1.213$, $p=0.299$). Likewise, both N2 and N5 (based on the maximum representation) from *the first run* showed a significant effect of the task type on the right posterior ($F(1,9)=6.563$, $p=0.031$), whilst

those from *the second run* showed no significant effect of the task type ($F(1,9)=0.001$, $p=0.975$). The results are summarised in Figure 7.25. As can be seen, there were significant increases of these discussed values during the experimental session from the first run.

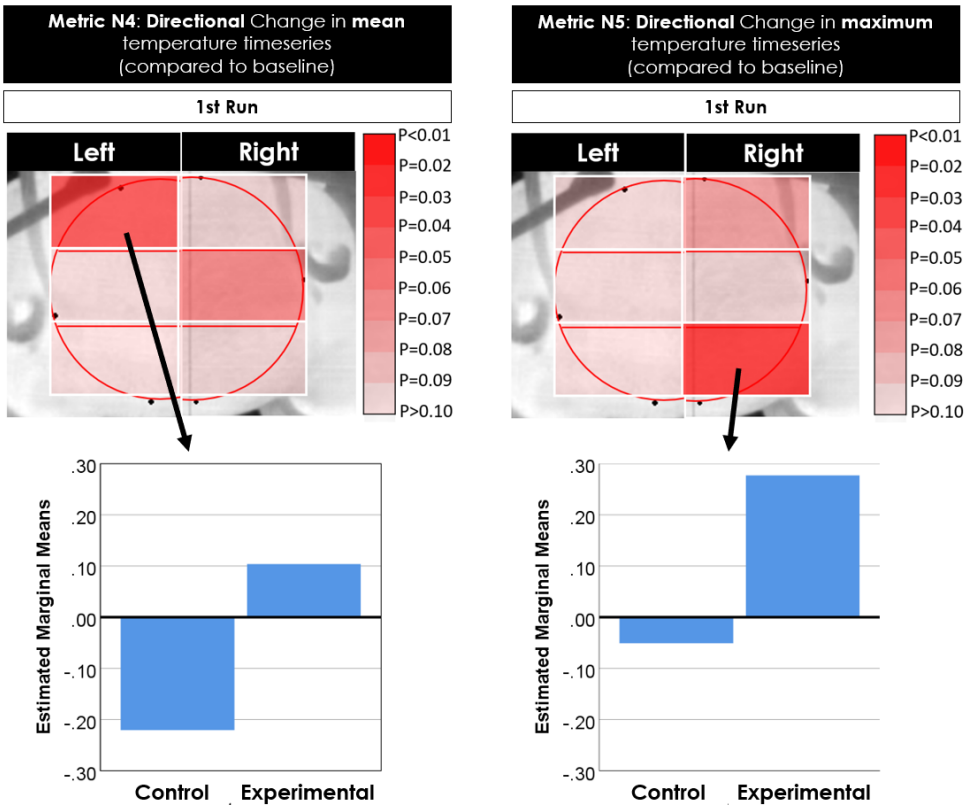


Figure 7.25. Results from significance tests of the task type **only** from **the first run** on basic metrics **N4, N5** (Directional change in mean/maximum temperature timeseries -compared to baseline) are visualised on each thermal image of the scalp over the left and right hemispheres. The left bar plot shows the N4 values from the left anterior area across each condition. The right bar plot shows the N5 values from the right posterior area across each condition.

7.4. Discussion

Inspired by Lombard’s earlier investigation on the scalp temperature associated with cognitive and emotional functions in 1878, we have explored the use of low-cost thermal cameras to capture cortical thermal signatures in relation to mental stress. To the best of our knowledge, this is the first investigation of the scalp temperature using thermal imaging. Accordingly, we have mainly contributed i) a novel contactless way to study stress-related cortical activity and ii) computational methods to capture cortical thermal signature and its variability.

Mobile thermal imaging enabling studies on cortical thermal signatures

From the neuroscientific, neurological literature exploring the brain of rats, we have found two interesting phenomena: i) regional CBF activation/deactivation pattern could be correlated with brain core temperature (Li *et al.*, 2017); and ii) brain core temperature affects cortical temperature and temporalis muscle (Busto *et al.*, 1987). Given this, we have assumed that it is possible to relate regional CBF to the scalp, cortical temperature although the phenomena were found from rats. With this starting point, we have explored temperature patterns on the scalp local areas above documented CBF activated regions in association with mental stress. Amongst documented regions, the most frequently reported main “stress” regions are: *right PreFrontal Cortex (PFC)* (Wang *et al.*, 2005; Causse *et al.*, 2017; Rosenbaum *et al.*, 2018), *left inferior frontal gyrus* (Taylor *et al.*, 1997; Soufer *et al.*, 1998; Pruessner *et al.*, 2008) and local regions in occipital/parietal lobes (Taylor *et al.*, 1997; Pruessner *et al.*, 2008).

In exploring thermal signatures from such regions, we must confront installation issues with thermometry (such as mercury thermometry used in Lombard, 1878 or modern resistance-based type) or heavy-weight thermal imaging systems which have been dominantly used in recent thermal imaging studies (as discussed throughout this thesis, they have focused on facial temperatures). Given this, we have taken advantage of the portability and lightweight of mobile thermal imaging to avoid any installation issues, in turn proposing novel recording set-ups, as shown in Figure 7.6 and Figure 7.13. The setups do not require heavy equipment or large spaces which are considered as main limitations of existing brain imaging methods (Boto *et al.* 2018). The proposed setups have allowed us to easily conduct studies using existing protocols such as mental arithmetic-based stress induction studies.

Importance of capturing variability of cortical thermal signature in assessing mental stress

In two separate studies (Study I: in unconstrained settings as an initial investigation in Section 7.2; Study II: in highly constrained settings to reliably obtain time varying cortical thermal signals in Section 7.3), firstly, we have explored the metrics capturing thermal directional information (metrics in Table 7.1 and 7.2). From both Study I and the first run in Study II, the scalp area above the occipital lobe (one of the main *stress* activated regions) produced a significantly higher amount of heat (increase) under the stress condition than the control, as

shown in Figure 7.12 Middle and Figure 7.25 Right. However, the phenomenon was not replicated in the second run of Study II. This may indicate that measuring directional changes in temperatures (increase in N5 in Figure 7.25 Right) is not a good way to capture cortical thermal signature. By contrast, the metric V2 - which is designed to improve the capturing of variance levels (in Table 7.3) - was able to capture highly varying signal patterns on the area during the experimental session than the control condition from both runs in Study II (Figure 7.24 Right). This can lead to a discussion of the important role of variability in assessing mental stress as in Chapter 6.

Furthermore, by contrast with Study I where we have focused on thermal directionality, the use of variability-driven metrics V1 and V2 in Study II has shown significant thermal activations on the right anterior area which the right PFC is situated underneath. As more recent studies including the one conducted in real world contexts beyond laboratory mental arithmetic (e.g. a stressful aviation task using a flight simulator in Causse *et al.* 2017) have repeatedly reported the significant activation of the right PFC (particularly, dorsolateral) in response to mental stressors (Wang *et al.*, 2005; Al-Shargie *et al.*, 2016; Causse *et al.*, 2017; Rosenbaum *et al.*, 2018), the statistically stronger activation we found from the area is very encouraging. However, V1 has also shown significant thermal activations on the right middle area. This is not supported by the literature exploring stress-induced CBF activation regions. The reason could be two fold: i) temperature on the area may be affected by the strong activation of the right PFC, or ii) the superior temporal gyrus (in Figure 7.5), which has been reported as a stress-induced CBF *deactivation* region (Taylor *et al.*, 1997; Pruessner *et al.*, 2008), may affect thermal variability on the area.

Amongst variability-related metrics (V1-V6), only skewness (spatially applied to a local ROI) based measures V1 and V2 have shown significant stress-induced cortical thermal activations. As we have assumed earlier, this result confirms that the spatial skewness measurement can be a good way to represent and track spatially differentiating patterns related to CBF-induced thermal propagation. This is interesting because spatial skewness has not yet been used to spatially represent two-dimensional temperature elements on a thermal image, whilst averaging has been the most widely used method in the literature (exploring facial thermal images) (Murthy & Pavlidis, 2006; Fei & Pavlidis, 2010; Abbas *et al.*, 2011; Lewis *et al.*, 2011; Gane *et al.*, 2011; Pavlidis *et al.*, 2012; Pereira *et al.*, 2015; Abdelrahman *et al.*, 2017). As we addressed in Chapter 5 (robust respiration tracking), the main limitation is that the averaging method is

prone to ignore small spatial changes and we have found that the use of skewness measurement can address the issue.

Limitations and future opportunities

Despite promising results and greater potentials of this work, there are limitations. Firstly, there is still a need for more refined registration techniques. Even with a controlled setup in Section 7.3, the orientation, size and shape of the head can affect the registration of local ROIs. If this is addressed, a higher spatial resolution (segmentations) of cortical measurement (than the current six segmented areas in Study II) can possibly be used, compared to the state-of-the-art EEG and fNIRS devices. One way to possibly contribute towards a better registration is to virtually project the brain on a scalp thermal image as illustrated in Figure 7.26. Secondly, participants' hair type is an issue to keep this approach from being adapted to more general cases since a thermal imaging device has difficulty in seeing through hairs or other organic materials.

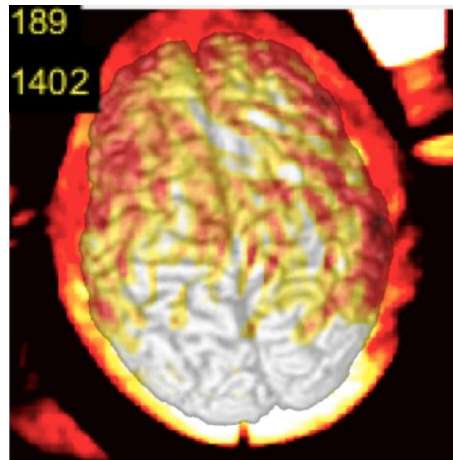


Figure 7.26. Visualisation of a *virtual* (mock-up) cortical mapping on a thermal image of the scalp to help a better registration.

According to Boto *et al.* (2018), there is a need for new techniques which can be more accessible to the general public than existing equipment for cortical activity measurements. Once the discussed limitations are resolved, this new mobile thermal imaging-based approach could provide new opportunities for a variety of real-world applications. As a next step towards this, we plan to build an open source platform for expanding this initial work and contributing to the research community in future work with such software libraries and data sets.

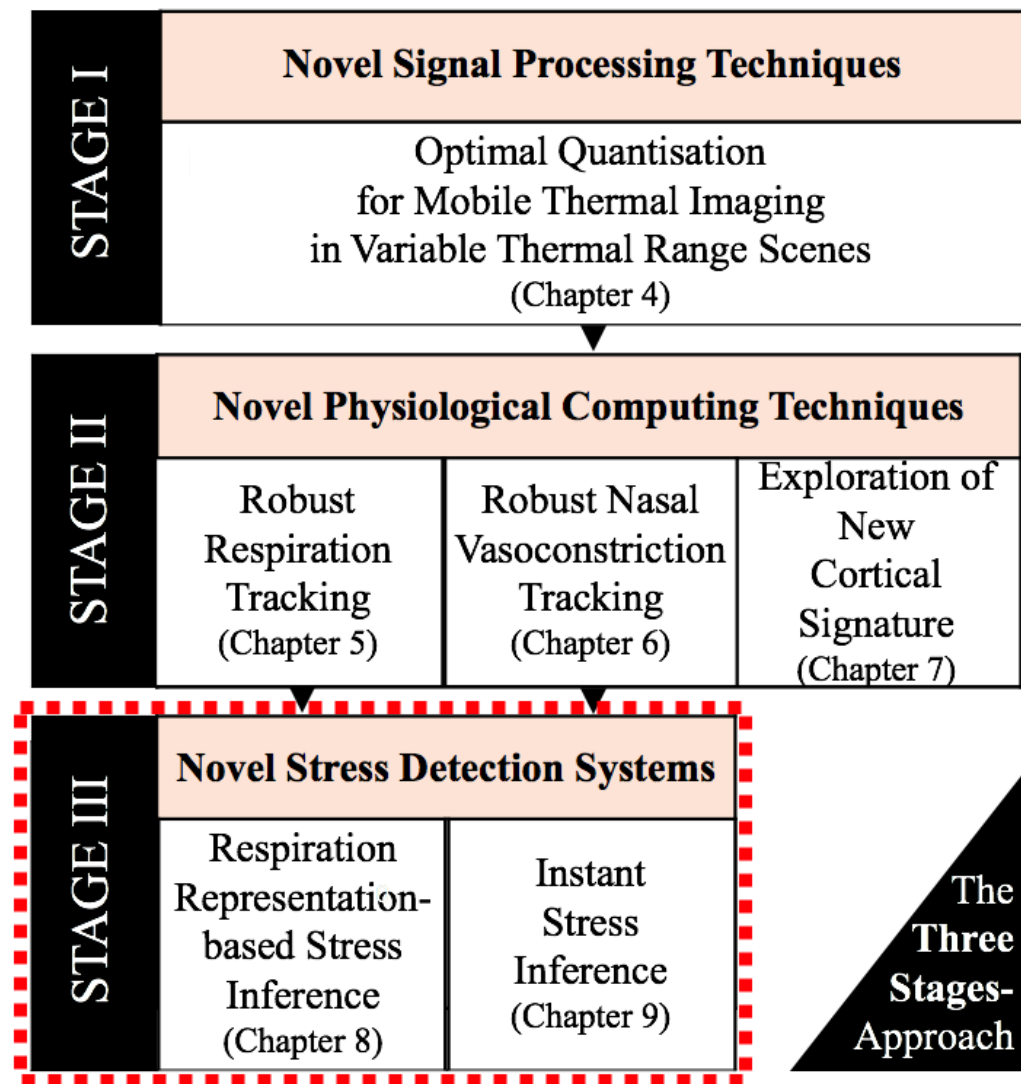
7.5. Chapter Summary

Based upon neuroscientific findings in the literature, we have undertaken a novel investigation on cortical thermal signatures using thermal imaging. Given the strong relationship between regional CBF activation and temperature propagation, we have proposed novel framework, computational methods and metrics to capture stress-induced cortical thermal variations. From our two separate studies conducted in both constrained and unconstrained settings, we have demonstrated that the key documented CBF stress activation regions can be associated with temperature. We have also identified the issue of head movement and hair type which interfere with capturing cortical thermal signatures. Finally, we have shown the importance of capturing variability of thermal signature (in this chapter, cortical thermal signature) in assessing mental stress.

Unlike respiratory (in Chapter 5) and vasoconstriction/dilation-related cardiovascular thermal signatures (in Chapter 6), which has achieved robustness in monitoring, this very initial finding is less likely to contribute to building automated stress recognition systems due to a lack of algorithms needed to automate the procedure of signature extraction. Hence, on the rest of this thesis we will focus on the more investigated signatures, respiratory and vasoconstriction/dilation-related cardiovascular signatures. Nonetheless, these findings in this chapter have shown positive possibilities of building a novel low-cost, mobile, non-contact way to observe cortical activities which can support a wide range of real-world contexts.

STAGE III

Novel Automatic Mental Stress Recognition Systems



Chapter 8

Automated Mental Stress Recognition using Respiratory Signature

In Stage III, we aim to contribute novel methods that can *automatically* recognise a person's mental stress based on mobile thermography-driven physiological computing, which has been proposed in previous chapters. Rather than just combining multiple signatures together to obtain better performance, we explore each signature in depth and its contribution to mental stress inference since their roles have shown to be weak in the computing literature (Engert *et al.*, 2014; McDuff *et al.*, 2016). In this chapter, we focus on respiratory thermal signature. Although it has been shown that stress significantly affects breathing patterns in the psychological, physiological literature (Grossman, 1983), existing computational methods have often failed to use breathing information in reliably assessing mental stress (McDuff *et al.*, 2016). To bridge this gap, we aim to build a novel technique to improve capturing of stress-related breathing information and achieve state-of-the-art performance in detecting mental stress levels.

In Section 8.1, we review existing techniques to automatically detect mental stress. In Section 8.2 we propose our computational approach which revisits the ability of breathing information in inferring mental stress levels. In Sections 8.3 and 8.4 we conducted a data collection study, dataset and evaluation methods. Section 8.5 proposes a labelling technique based on perceived mental stress levels, which is followed by reporting results in Section 8.6. This chapter ends with a discussion in Section 8.7 and a chapter summary in Section 8.8.

8.1. Background and Research Question

Breathing is an important vital process controlled by the Autonomic Nervous System (ANS). The monitoring of breathing patterns can be informative of a person's mental and physical condition. Given the fact that psychological stress affects the physiological processes mediating human vital signs (Grossman, 1983; Brosschot *et al.*, 2006; Nash & Thebarg, 2006), researchers have investigated the possibility of using breathing signals together with other physiological signatures, for instance, from cardiovascular activity (heart rate, heart rate variability, blood pressure), perspiratory activity (e.g. skin conductance level) and neural activity (e.g. alpha, beta waves from EEG), to automatically assess people's stress levels (e.g. Healey and Picard, 2005; Hernandez et al., 2011; Hong et al., 2012; Hosseini and Khalilzadeh, 2010; McDuff et al., 2016; Puri et al., 2005).

Minor role of breathing features in automatically detecting mental stress

The majority of approaches for building automatic stress recognition systems require already defined hand-engineered features representing physiological cues. In the case of respiratory activity, average breathing rate and tidal volume from a time window (e.g. 2 minutes) have been most widely used (Hong *et al.*, 2012; McDuff *et al.*, 2016). However, the results of studies using such features have shown weak contributions of breathing features compared with features from other physiological signals. For instance, McDuff *et al.* (2016) show the most informative feature is HRV with breathing rate playing a minor role in a binary stress classification task (in their case, resting and stressful sessions as binary classes) using multiple physiological features (i.e., heart rate, breathing rate and HRV), from which their computational model achieves 85% accuracy. This may indicate that breathing rate itself is not the best informative metric for respiration. Indeed, directional changes in respiration rate during stressful conditions have often been reported inconsistently. For instance, while Masaoka *et al.* (1997) reported an increase of a person's breathing rate in response to mental stressors, Hong *et al.* (2012) reported a significant drop of the rate along with stress.

Despite the minor role of breathing features in automatically assessing mental stress and the incongruent findings of breathing rate changes from the literature, it is inarguable to say that a person's respiration pattern is varied when experiencing stressful events (Grossman, 1983). In addition, when exploring HRV which is known to be informative to mental stress, multiple parameters are computed to capture its complex dynamics. This may indicate a need for exploring

the complexity of breathing patterns rather than just discrete rate. Furthermore, as respiratory activity is often coupled with cardiac activity (for instance, respiration rate can be roughly estimated from blood volume pulse measurement, e.g. photoplethysmography, Gastel *et al.*, 2016), even only the use of the information for breathing dynamics may contribute to automated mental stress recognition as much as HRV does.

This leads to our key research question: *How can respiratory thermal signature itself contribute to building automatic stress recognition systems?* Accordingly, in this chapter, we explore how to capture and represent complexity of breathing dynamic patterns (i.e. variability rather than just rate) to improve its ability in inferring a person's stress levels. From this, we can take newly represented breathing information as low level features from which advanced machine learning techniques learn informative high level features.

Automatic feature learning techniques

Automatic feature learning is a power centralised in most successfully and widely used deep learning approaches in computer vision and pattern recognition (LeCun *et al.*, 2015). This helps automatically find good features during the machine learning process. As even carefully hand engineered-feature extractors could fail to generalise to unseen data sets (LeCun *et al.*, 2015), this could be a potential solution to improve our understandings of physiological and behavioural patterns in relation to a person's affective states.

Indeed, feature learning techniques have been explored to detect facial and voice affective expressions. For instance, Liu *et al.* (2014) used the deep belief network for better characterisation of features in relation to a certain type of facial expressions. Lane *et al.* (2015) proposed a model which could analyse multiple large scale speech datasets using a coupled deep neural network for automated stress detection. These approaches have achieved greater performance in classification tasks, in comparison with non-automatic feature learning classifiers (e.g., SVM, Gaussian mixture models).

More recently, researchers have shown the capability of Convolutional Neural Network (CNN) based models (LeCun & Bengio, 1995; Krizhevsky *et al.*, 2012; Lin *et al.*, 2013; Szegedy *et al.*, 2015) in a variety of problems spanning from image classification to semantic segmentation, demonstrating that the models themselves can produce better learning performance than earlier learning approaches. The recent accomplishments in CNNs have been adopted to improve affect recognition performance: for instance, arousal and valence prediction in audio (Trigeorgis *et al.*,

2016), facial action unit identification (Ghosh *et al.*, 2015), and facial expression recognition in video (Yu & Zhang, 2015), but not for respiratory signals. Accordingly, this chapter explores the use of CNNs for automatically learning more informative features from condensed breathing information, in turn revisiting roles of breathing signals in assessing mental stress and contributing towards state-of-the-art stress recognition systems.

8.2. Proposed Computational Methods

In this section we propose 1) a novel representation of breathing pattern dynamics; 2) an automatic feature learning framework which does not require feature hand-crafting. Although the proposed approach for automatic stress detection is sensor-independent, this is implemented on mobile thermography-based respiration tracking proposed in Chapter 5.

8.2.1. Novel Representation of Respiration Variability

Figure 8.1 illustrates the process for collecting breathing patterns using the low-cost mobile thermal imaging camera (dimension: 18mm x 26mm x 72mm). Thermal videos are captured using the camera attached to a smartphone placed at a maximum distance of 55cm from a person's nostril area (i.e., our ROI) as shown in figure 8.1a. We use our respiration tracking algorithm proposed in Chapter 5 to recover one-dimensional breathing signals from the nostril ROI on the thermal videos as shown in figure 8.1b.

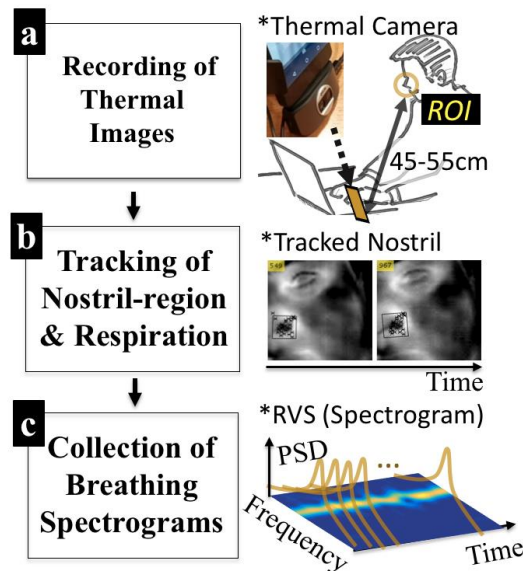


Figure 8.1. Process for the respiration variability spectrogram collection through thermal imaging.

Figure 8.1c shows a two-dimensional spectrogram automatically transformed from one-dimensional signal sequences. This is for condensing dynamic time-frequency domain information (i.e., respiration variability) into an image (we call this Respiratory Variability Spectrogram, RVS, afterwards). More specifically, a two-dimensional spectrogram can be constructed by stacking a Power Spectral Density (PSD) vector of short-time-window respiration signals (i.e., Equation (5.7) in Chapter 5) over time until the end of recordings. The PSD function handles the short-time autocorrelation that identifies similarities between neighbouring signal patterns, being of use to examine respiration variations in a short period. The RVS can be expressed as:

$$RVS(f, \hat{t}) = \sum_k R_{ww}(k, \hat{t}) e^{-j2\pi f k} \quad (8.1)$$

where $R_{ww}(k, \hat{t})$ is the short-time autocorrelation output of windowed breathing signals from Equation (5.6) in Chapter 5, k is a time lag to examine the similarity, \hat{t} is the discrete time information (e.g., $\hat{t} \in [1s, 100s]$), and f is the frequency domain information.

For the implementation, as used in Chapter 5, the normal breathing frequency range of healthy adults, $[f_1=0.1\text{Hz}, f_2=0.85\text{Hz}]$, was used to band-pass filter a one-dimensional breathing signal in Equation (8.1) for reducing magnitudes of frequencies beyond the range of interest. f in Equation (8.1) has the same range scale. The time length of the short-time window is set to 20 seconds (as $\hat{t}_{\max} = 10s$ in Equation (5.6)) and moves forwards in 1 second intervals. Lastly, so as to produce an integer RVS matrix (i.e., to be easily taken as a grey-image), we converted the frequency in $[f_1, f_2]$ into an integer in a new scale $[1, m]$ (here, $m=120$) using a linear transformation $y = T(f)$. Note that this new bi-dimensional input signature can be applicable to any type of breathing measurement (e.g., chest-belt, expensive high-resolution thermal camera), despite our focus upon low-cost low-definition thermal imaging.

Examples of the RVSs (collected during our stress-induction study in the next section) are shown in Figure 8.2. As discussed in the literature (Hernandez *et al.*, 2011), physiological signals can vary strongly between people; for example, S4 shows a particularly different spectrogram from the others in Figure 8.2a. Although the breathing dynamics (i.e., variation patterns) are different between every individual, the higher the levels of stress a person experienced, the more irregular the patterns appear to be.

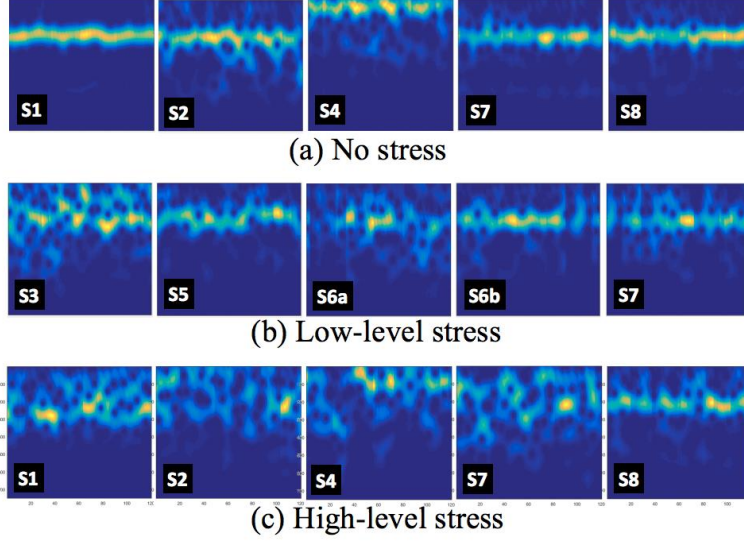


Figure 8.2. Examples of respiration spectrograms clustered according to participants' self-reported stress scores: (a) No-stress, (b) Low-level stress, (c) High-level stress.

8.2.2. Automatic Feature Learning using Convolutional Neural Network (CNN)

In comparison with other physiological signatures (e.g. heart rate, EEG), breathing patterns usually have a narrow and low frequency bandwidth (e.g., between 0.1Hz and 0.85Hz). In other words, it requires relatively long-term measurements to collect a sufficient amount of data which can be used in an automatic feature learning process. One possibility is to use a pre-trained network which is trained on a larger dataset, referred to as an efficient way to learn about small datasets (LeCun *et al.*, 2015). However, low-level features learnt from other domains, such as ImageNet data (Krizhevsky *et al.*, 2012), may not fit dynamic breathing patterns. Inspired by basic transformation-based data augmentation methods in deep learning frameworks (e.g. cropping, zoom-in) (Krizhevsky *et al.*, 2012), we propose a unidirectional sliding cropper with a square window to augment the RVS dataset while preserving each label, defined as:

$$I_i(x, y) = RVS(T^{-1}(y), i + x), \quad x \in \{1, 2, \dots, m\} \quad (8.2)$$

where $i \in \{1, 2, \dots, \hat{f}_{\max} - m + 1\}$ is in an ascending order (1 second interval), m is the length of each unit image, y in Equation (8.2) is inversely transformed (T^{-1}) to f (in Equation (8.1)), in turn producing $(\hat{f}_{\max} - m + 1)$ images of size $m \times m \times 1$ (height x width x channel). With m set to 120 in the subsection above, each unit spectrogram image has 120x120 size as visualised in Figure 8.3.

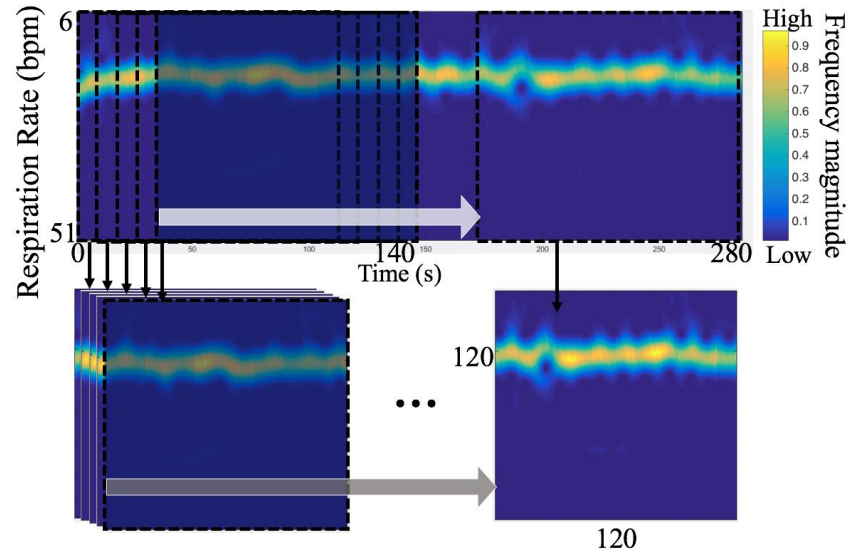


Figure 8.3. Proposed data augmentation technique using a uni-directional sliding cropper with a 120x120 square window. The colour scale indicates the frequency magnitude (i.e. PSD).

The CNN network in Figure 8.4 consists of two convolutional layers, two pooling layers, and one fully connected layer. Although additional modifications can be applied to this basic form of CNN (LeCun & Bengio, 1995), our aim in this chapter is to investigate the possibility of using automatic feature learning to improve assessment of stress levels from physiological patterns. In addition, we aim to build a low computational cost deep learning network for running on embedded or mobile systems, as seen in Figure 8.1a. Each image patch of 120 x 120 is resized to 28 x 28 using a basic bi-cubic interpolation to be fed forward to the first convolutional layer. The size of the resized patch was chosen to use the CNN structure proposed in (LeCun & Bengio, 1995). The first layer filters the resized input image with n kernels (this value is set according to the number of levels to classify in Section 8.4) of size 5 x 5, the second convolutional layer has j kernels of size 5 x 5 and each pooling layer has averaging filters of size 2 x 2 applied with a stride of 2. A sigmoid function, widely used in artificial neural networks, is connected to every convolutional and fully connected layer to induce non-linearity. In the case where a large dataset is available, a Rectified Linear Unit (ReLU) can replace this to boost the learning speed (Krizhevsky *et al.*, 2012). The output of the final pooling layer is fully connected to each output neuron which corresponds to each targeted class. Stochastic gradient descent (SGD) is applied to supervised fine-tuning during back propagation to reflect classification errors on training data sets.

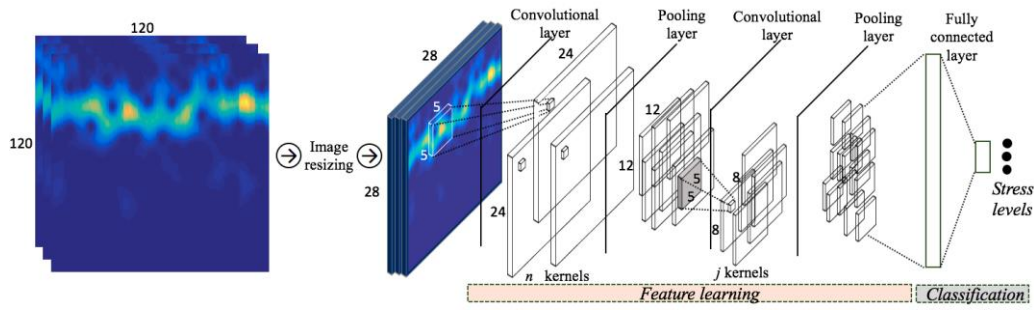


Figure 8.4. The mainly used CNN architecture consisting of two convolutional layers, two pooling layers and one fully connected layer.

8.3. Experimental Protocol and Dataset (*DeepBreath*)

As discussed in Section 3.3, two widely used tasks for inducing mental stress were selected for our purpose: the Stroop Colour Word test (Stroop, 1935) (denoted as *Stroop*) and the mathematics test (denoted as *Math*) used in Pruessner *et al.* (2008), Hong *et al.* (2012) and McDuff *et al.* (2016). Each task has both an easy and a hard session. This is to ensure a good spread of induced stress levels within each task. In the Stroop task, all participants were required to name the colour of a word. In the easy session of the first task, the meaning of a word and its font colour were congruent (e.g. the word *red* written in *red*). In the hard session, they were incongruent (e.g. the word *yellow*, but written in *red*). The Math task requires the participants to repeatedly subtract (mentally) a certain number (e.g. 1, 13) from a four-digit number (e.g. 5000), as used in Chapters 6 and 7. It is expected that difficult sessions in Math and Stroop tests would lead to higher stress levels than the easy sessions.

8 healthy adults (3 females) (aged 18-53 years, $M=30.75$, $SD=10.22$) were invited from the subject pool service of University College London. Each participant was given the information sheet and provided informed consent prior to data acquisition. The study took place in a quiet lab room with no distractions (and no room temperature control). During each Stroop and Math task, each participant answered the task questions using a mouse on a laptop. In the Stroop task condition, participants had to click-select the right answer amongst different colour options while pronouncing the colour aloud. Every question was shown for 1.5 seconds as in (Åkerstedt *et al.*, 1983). In the Math task condition, each participant typed an answer using a mouse on a virtual keypad (i.e. GUI in the screen). Each arithmetic question was shown for 7.5

seconds which was set based on a pilot study. After each answer, participants received sound feedback that informs them whether the answer was correct or not.

The flow for our experimental design is shown in Figure 8.5. After a 5-minute relaxation period, each participant was asked to go through the two types of cognitively demanding tasks (i.e., *Stroop*, *Math*). Before and after each demanding task (i.e., during *Q*), all participants were instructed to rate the subjective feeling based on a continuous 10-cm Visual Analogue Scale (VAS). VAS is a standard approach used to avoid the non-parametric property of the Likert scale (Bijur *et al.*, 2001). In this study, the main question was: “*Did you feel mentally stressed?*” (ranging from 0cm, not at all, to 10cm, very much). The easy and difficult sessions of each task type (i.e., *Stroop* 1,2 and *Math* 1,2) were counter-balanced.

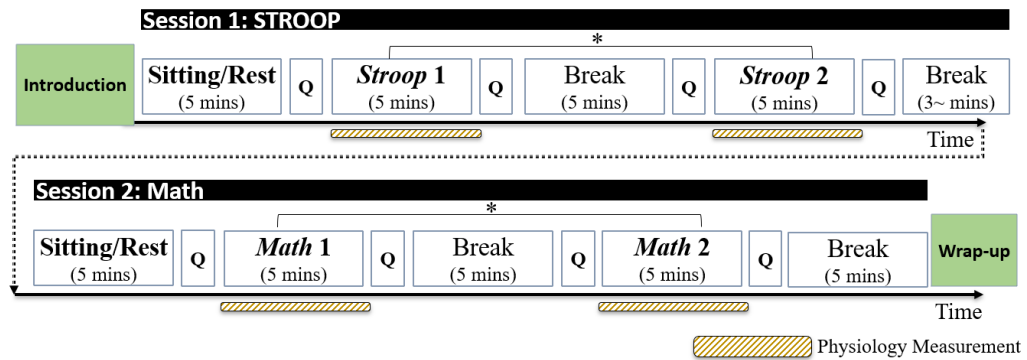


Figure 8.5. Flow chart of the proposed mental workload/stress induction experimental protocol: the first session is Stroop Colour Word task and the second session is the arithmetic solving task. *Tasks in each session were counterbalanced in Latin squared design.

The face of each participant was recorded using a low-cost thermal camera (as in Figure 3.2 in Section 3.2) only during Stroop and Math tasks in Figure 8.5. The whole sessions took 63 minutes - 72 minutes per each participant. Using the techniques proposed in Section 8.2, each video segment was transformed into an augmented set of spectrograms with the associated self-reported stress level (i.e. the scores reported at the end of each session). This process produced 3936 augmented input images (*DeepBreath* dataset) of size 120 x 120 (i.e. an average of 492 (SD=91.49) for each 8 participant) which are publicly available at <http://youngjuncho.com>.

8.4. Evaluation Method

First, we tested significance on subjective stress self-reports (VAS-based scores) across two different stress induction tasks. This is to mainly check whether four tasks (easy and hard sessions of the Stroop and Mental Arithmetic) induce different stress levels. With one Independent Variable (IV): a task type with four conditions (along with the four tasks), we conducted the normality test using the Shapiro-Wilk method with the aim to see if the scores are normally distributed for each level of the IV. As one set of data was skewed, we used Friedman’s analysis of variance (ANOVA). For the post hoc analysis of paired conditions, we conducted the nonparametric Wilcoxon signed rank test.

Second, we conducted two separate classification tasks: three-classes and binary tasks for evaluating stress detection performance of the proposed methods. The three-class classification task is to test the capability in discriminating multiple stress levels. As shown in Figure 8.2, breathing patterns labelled as high and low-level stress share a certain level of similarities. Given this, we also undertook a binary classification problem by combining the low-and high-level stress into a *stress* class to discriminate events of *stress* from *no-stress*. As for parameters of the proposed methods, the number (n) of kernels (i.e., filters) of the first convolutional layer was set to 12 for the three-class classification (i.e., three levels of stress classification) and 9 for the binary classification. The number (j) of kernels of the second convolutional layer was set to 24 and 18, respectively. In a training process, a fixed learning rate of 0.5 and a batch size of 50 were used for 300 epochs.

Furthermore, we implemented three single hidden-layer neural networks as basic automatic feature learning methods to compare performance of the proposed *CNN-based* model. The first single hidden-layer neural network (denoted as *NN1*) uses the resized image patches (i.e., 28x28 in the proposed model in Figure 8.4) as input and includes one hidden layer of node size 30. The second shallow network (denoted as *NN2*) uses half-sized image patches (60x60) and a hidden layer of size 45. The last network (denoted as *NN3*) uses the original images (120x120) and a hidden layer of size 300.

We trained and tested each model using a k-fold leave-one-*subject*-out (LOSO) cross-validation. This cross validation method has been widely used to test the ability to generalize to unseen participants’ physiological data (Hernandez et al., 2011). As illustrated in Figure 8.6, at every fold, all data sequences from the participants except for one were used to train the networks, and the data from the left-out participant was used for testing.

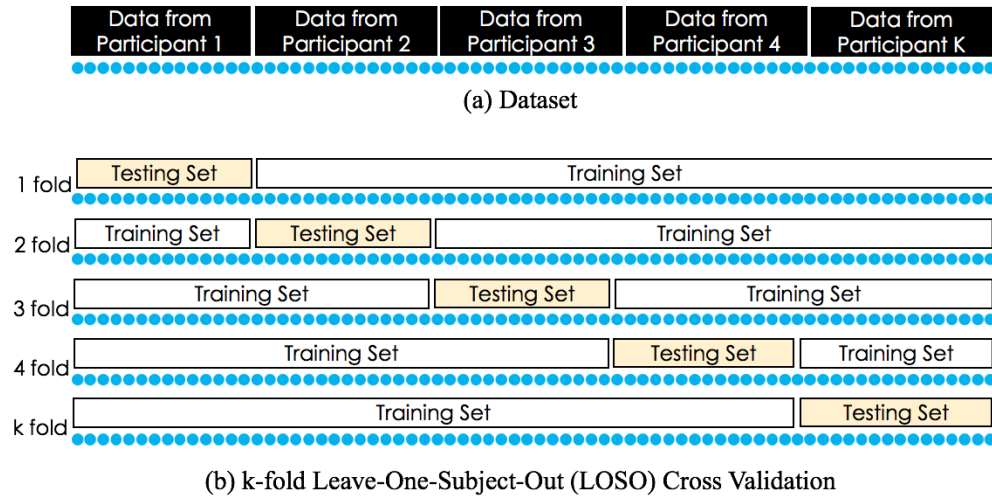


Figure 8.6. k-fold leave-one-subject-out (LOSO) cross validation method: (a) dataset collected from k participants, (b) splitting the dataset into training and testing sets for each fold.

8.5. Perceived Mental Stress Levels and Labelling

The boxplot in Figure 8.7 shows the distribution of the self-reported scores over the four tasks: StroopEasy $M=2.00$, $SD=2.07$; MathEasy $M=2.00$, $SD=2.27$; StroopHard $M=4.18$, $SD=2.75$; MathHard $M=4.75$, $SD=2.87$. As data for the easy arithmetic task were not normally distributed violating the normality assumption of repeated measures ANOVA (Shapiro-Wilk: $p=0.033$), we conducted the Friedman test to determine whether there was a statistical mean difference in the perceived mental stress levels across the four tasks. The results showed a significant effect of the task type on the perceived stress levels ($\chi^2(3)=9.554$, $p=0.023$). From the post-hoc analysis using the Wilcoxon signed rank test with an alpha value of 0.008 ($\approx 0.05/k$ where $k=6$, the number of pair-wise comparisons, which can be calculated by $n(n-1)/2$ where n is the number of conditions), perceived stress scores were not significantly different across tasks (e.g. StroopHard - StroopEasy: $Z=-1.577$, $p=0.115$; MathHard - MathEasy: $Z=-2.41$, $p=0.016$; MathHard - StroopEasy: $Z=-2.136$, $p=0.033$). This makes a difficulty in using the task type as a label for stress classification tasks. This leads us to developing a new labelling strategy based on perceived mental stress scores.

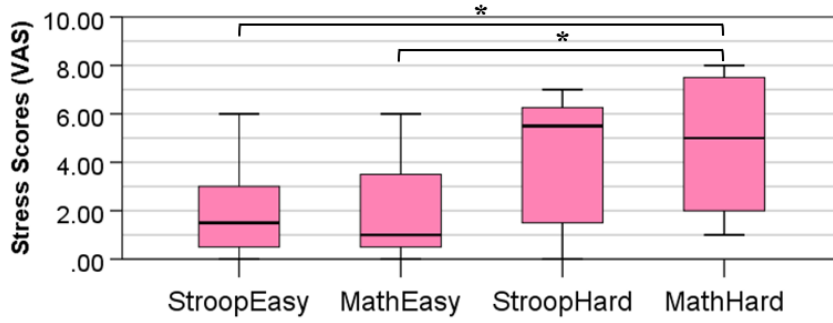


Figure 8.7. Boxplots (95% confidence interval) of self-reported stress scores from the 8 participants across each task (StroopEasy, MathEasy, StroopHard, MathHard).

First, we normalised participants' stress ratings to reflect inter-person variability in ratings (Hernandez *et al.*, 2011). For each participant, the normalisation was based on feature scaling using his/her minimum and maximum scores. Second, we used k-mean algorithm to cluster the normalised scores into the three stress classes. We used *k*-means because it had been shown to be effective in dealing with self-reported ratings (Salmivalli *et al.*, 1999; Kjeldstadli *et al.*, 2006). The clusters were then used to label the breathing spectrogram dataset, i.e., each spectrogram was labelled using the cluster the person's self-report belongs to. Figure 8.8 summarises the results. As intended, the distribution of normalised self-reports (Figure 8.8a) shows different levels of perceived stress across each task session. Figure 8.8b shows the clustering of the normalised subjective ratings into the stress classes. This new label technology is explored more in-depth in Chapter 9.

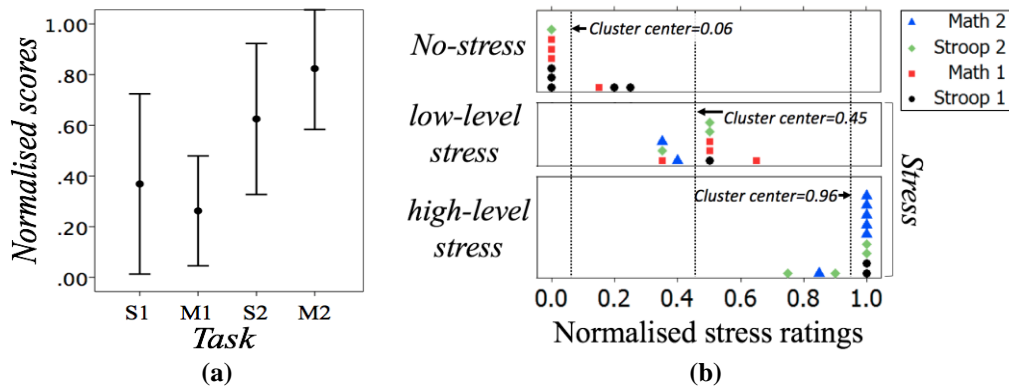


Figure 8.8. Normalised subjective VAS scores: (a) distribution of the self-report per participant (95% confidence interval), (b) clustered results using k-means. Colours and shapes of each point indicate the task type and difficulty levels. (Easy: Stroop 1(S1) and Math 1(M1); Difficult: Stroop 2(S2), Math 2(M2)).

Here we note that we used the data collected from the four task sessions (i.e., S1: easy Stroop, M1: easy Math, S2: difficult Stroop, M2: difficult Math), however not the ones from the initial relax period (widely used as baseline). There are rationales for this: first, the tasks effectively induced the different stress levels from no stress to the high level of stress given that there were significant differences between each pair of clustered groups (all t-test comparisons had $p < 0.001$; *none-low*: $t(18)=8.939$, *low-high*: $t(20)=13.390$, *none-high*: $t(20)=23.651$). In addition, there were no significant differences between subjective ratings collected from the initial relax period and those clustered into *no-stress* group ($t(16)=1.211$, $p=0.244$), suggesting the no-stress ratings to be a good baseline. Second, the recognition of non-stress in the relax periods would have been easier because of the lack of verbalisation and movement (e.g., mouse clicking) which could all affect the breathing patterns (independently of stress levels) (Hong *et al.*, 2012). Indeed, the tasks including the easy and difficult sessions had been purposely designed to verify the robustness of the approach to work in verbal contexts and unconstrained sedentary movement settings and the use of the relax periods as baseline would have reduced the validity of the evaluation.

8.6. Results

The dataset is composed of the augmented breathing spectrogram sequences of size 120 (i.e. frequency range) x 120 (window length) x 3936 (the number of two-dimensional spectrograms) and the stress labels (i.e. either with binary or three-class labels). In the three-class classification task, data sets were almost balanced (No-stress: 1305, Low-stress: 1207, High-stress: 1424 instances). For the binary classification task, however, combining both low- and high-stress classes resulted in an uneven class distribution (Stress class: 2631 instances).

Three-class classification task

Figure 8.9 shows accumulated confusion matrices and accuracies from 8 LOSO fold testing sets for the classifiers CNN and NN1, which have the same input of size 28x28. Compared with the majority case (the ratio of the number of data in the majority class to the total number of dataset: $1424/3936=36.18\%$), both models achieved higher accuracies (CNN: 56.73%, NN1: 51.47%). the highest accuracy ($M=56.52\%$, $SD=17.58\%$) was achieved with the CNN model,

while $M=51.89\%$ ($SD=10.36\%$), $M=51.02\%$ ($SD= 8.33\%$), 53.65% ($SD=10.39\%$) were the results from NN1, NN2, NN3, respectively, as shown in Figure 8.10.

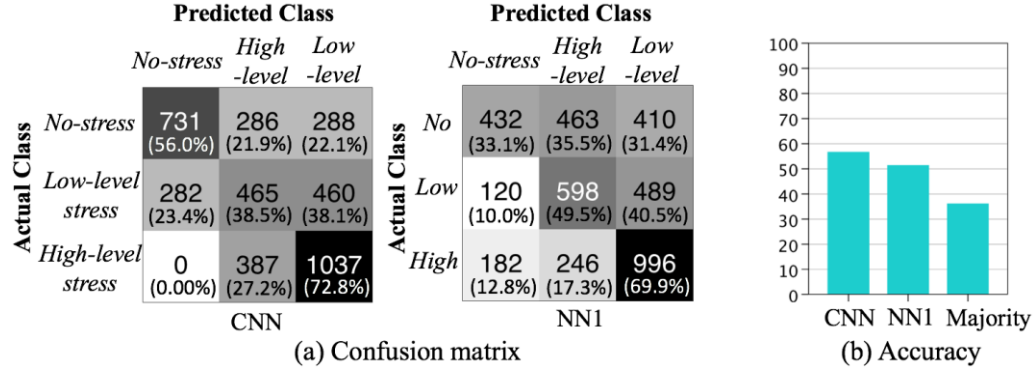


Figure 8.9. Test results from the three-class classification task (None, Low, High-level stress) with the CNN and the single hidden layer neural network NN1: (a) accumulated confusion matrix from 8 LOSO folds testing sets (each block consists of the sum and average over the LOSO cross validation results, the colour represents the number of each prediction (i.e., black: highest, white: lowest)), (b) accumulated accuracy from 8 LOSO fold testing sets.

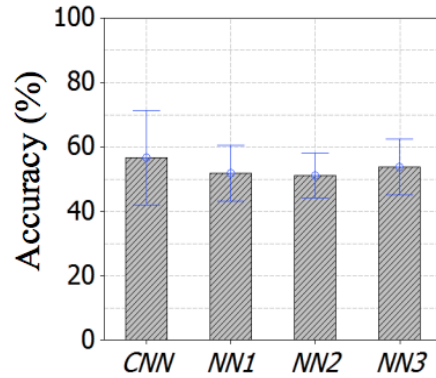


Figure 8.10. Summary of mean classification accuracy results across 8 LOSO fold testing sets for three-class classification (i.e., none, low, high-level stress). CNN: 5 layers CNN using low resolution RVS images (28x28), NN1: single hidden layer using low resolution RVS images (28x28), NN2: single hidden layer using medium resolution RVS images (60x60), NN3: single hidden layer using high resolution RVS images (120x120).

Binary classification task

Figure 8.11 shows accumulated confusion matrices and accuracies from 8 LOSO folds. Compared with the majority case (2631/3936=66.84%), both models achieved higher accuracies (CNN: 81.71%, NN1: 72.99%). The accuracy results for each LOSO fold confirmed that the CNN produced a much higher accuracy (M=84.59%, SD=19.34%) than the shallow networks (NN1: M=74.89%, SD= 19.12%, NN2: M=74.81%, SD=17.58%, NN3: M=77.31%, SD=19.45%), as shown in Figure 8.12. In comparison with NN1, the CNN performed significantly better in predicting both the *no-stress* (713 corrected predictions from CNN versus 467 from NN1) and *stress* states (2503 from CNN versus 2406 from NN1) as shown in Figure 8.11. Overall, the results showed the CNN-based automatic feature learning model outperformed in classifying the condensed breathing information on our proposed RVSs for each labelled group.

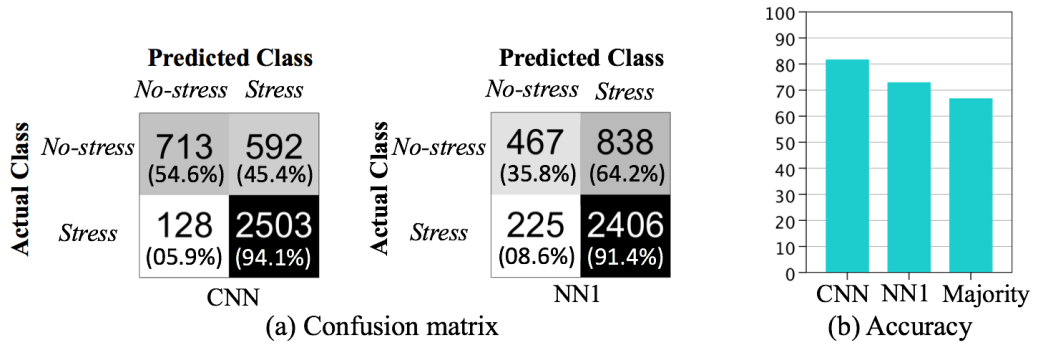


Figure 8.11. Results from the binary classification task (No-stress and stress) with the CNN and the single hidden layer neural network NN1: (a) accumulated confusion matrix from 8 LOSO fold testing sets (each block consists of the sum and average over the LOSO cross validation results, the colour represents the number of each prediction (i.e., black: highest, white: lowest)), (b) accumulated accuracy from 8 LOSO fold testing sets.

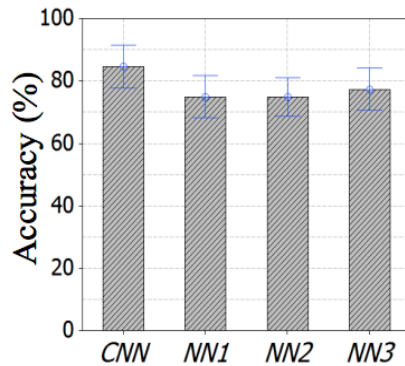


Figure 8.12. Summary of mean classification accuracy results across 8 LOSO fold testing sets for binary classification (i.e., no-stress, stress). CNN: 5 layers CNN using low resolution RVS images (28x28), NN1: single hidden layer using low resolution RVS images (28x28), NN2: single hidden layer using medium resolution RVS images (60x60), NN3: single hidden layer using high resolution RVS images (120x120).

8.7. Discussion

Contributing to the design of novel low-cost, non-contact thermography-based stress recognition systems, in this chapter we have focused on how to capture respiration variability concisely. Accordingly, we have contributed a new input signature: (i) respiration variation spectrogram condensing breathing dynamics, (ii) an automatic feature learning method supported by a deep learning framework, and (iii) a new thermal image dataset labelled with the stress levels from our structured multi-level stress induction tasks. Below are detailed discussions of these contributions.

The use of the respiratory signature alone contributes to the state-of-the-art stress inference performance

Given the fact that stress affects physiological processes (Grossman, 1983; Brosschot *et al.*, 2006; Nash & Thebarg, 2006), respiration amongst other physiological processes have been explored for automatic detection of stress (Healey & Picard, 2005; Hong *et al.*, 2012; McDuff *et al.*, 2016). Most of these works have focussed on gross statistical features. In particular, in the case of breathing, average breathing rate (over a specified window) has been mainly used, however its contribution to automatic stress recognition to date has been weak (Hong *et al.*, 2012; McDuff *et al.*, 2016). Inspired by Grossman (1983)’s note of breathing irregularity which could be induced by mental stress, we have gone deeper into stress-induced breathing dynamic patterns, proposing a new way to represent breathing variability. This representation technique is based on a bi-dimensional spectrogram of a uni-dimensional breathing signal to represent breathing variability in order to capture, in a compact way, breathing dynamics over time. The compactness of the representation enables the use of advanced machine learning techniques.

The accuracy of results from the first approach show that the proposed deep learning-based stress recognition model produces better performance (mean accuracy across 8 folds: 84.59% in the binary case, 56.52% in the multi-class case) than the best results from the shallow learning method (single-layer NN: *binary* 77.31%, *multi-class* 53.65%). However, these accuracy results were not significantly different. This may be related to the relatively small number of participants, although this number was set following existing work (Åkerstedt *et al.*, 1983). Regardless of this differentiation of deep and shallow networks made in this evaluation, this is the first study aimed at learning breathing patterns to automatically recognise a person’s mental stress level without hand-engineered feature extractors which require in-depth domain knowledge and highly matured skills. In particular, our proposed respiration variability spectrogram inputs (RVSs) are

able to capture the dynamic breathing information from a very short-lasting (e.g., a sudden drop and increase) to long-lasting respiration changes. The proposed convolutional kernels-based model can support automatic feature learning of the input of low-level features (RVSs), and in turn contribute towards state-of-the-art stress detection performance with the use of only breathing information.

Compared with the literature suggesting low discriminative power of breathing in discriminating between stress levels (McDuff *et al.*, 2016), the proposed model shows that respiration signals alone can lead to above chance levels in 3-classes recognition (i.e., high, low, no stress). In addition, in the case of binary recognition, our performance was similar to the one reported in the work achieving state-of-the-art performance (McDuff *et al.*, 2016) with the difference that multiple physiological channels were used rather than just breathing (McDuff *et al.*, 2016). We assume that our high performance is due to the use of breathing variability rather than just breathing rate. Hence, we expect the combination of breathing variability with other signals may lead to further improvement in the stress recognition performance. This should be tackled as future work.

It should be noted that despite the fact that the evaluation tasks did not intentionally require large movements, people were free to move as needed (e.g., inputting their responses, or displaying stress related expressions e.g. turning or bending the head to escape the source of stress and frustration, Kleinsmith & Bianchi-Berthouze, 2013). We expect that even larger movements and outdoor situations would be acceptable, as in Chapter 5 the nostril ROI tracking and respiration tracking were thoroughly tested in both highly mobile contexts and highly-variable thermal range scenes.

Generalisation of the proposed Representation-based Approach to Other Physiological Signatures

Despite the focus of this thesis upon thermal imaging, our approach to automatic stress detection is sensor-independent. In particular, the representation approach proposed to capture respiration variability can be generalised to other periodic physiological signals (e.g. blood volume pulse from PPG) as shown in Figure 8.13 - a spectrogram generation network is to help to generate variability spectrograms along with each type of physiological cues. The output of this network can be connected to an artificial neural network for automatic feature learning. Figure 8.14 shows examples of a pair of the RVS and blood volume pulse variability spectrogram

from a rest period and a stress induction task period (here, the Math hard). As the high performance was achieved from this approach, other types of physiological signals may benefit from this representation to produce more informative stress signatures (e.g. in Chapter 9, we show HRV metrics derived from blood volume pulse are poorly correlated with perceived mental scores). We believe this would be an interesting future direction in solving a wide range of affect recognition problems.

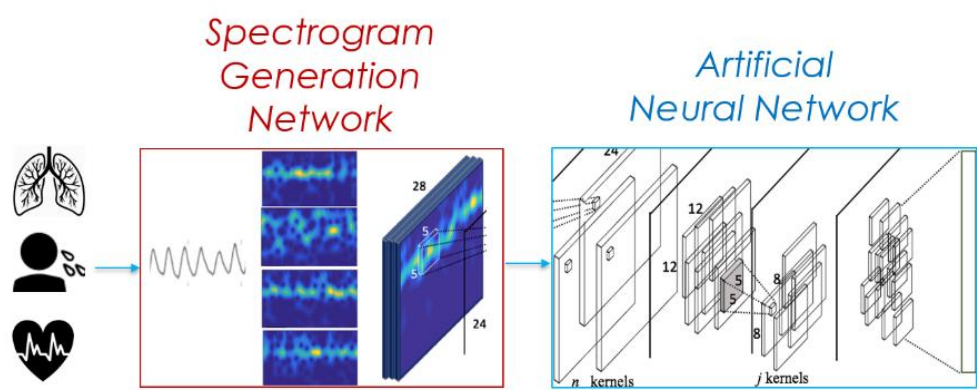


Figure 8.13. A generalised representation model for different types of physiological signals: a spectrogram generation network is to help generate variability spectrograms of each physiological pattern (e.g. blood volume pulse, etc). The output of this network can be connected to an artificial neural network for automatic feature learning.

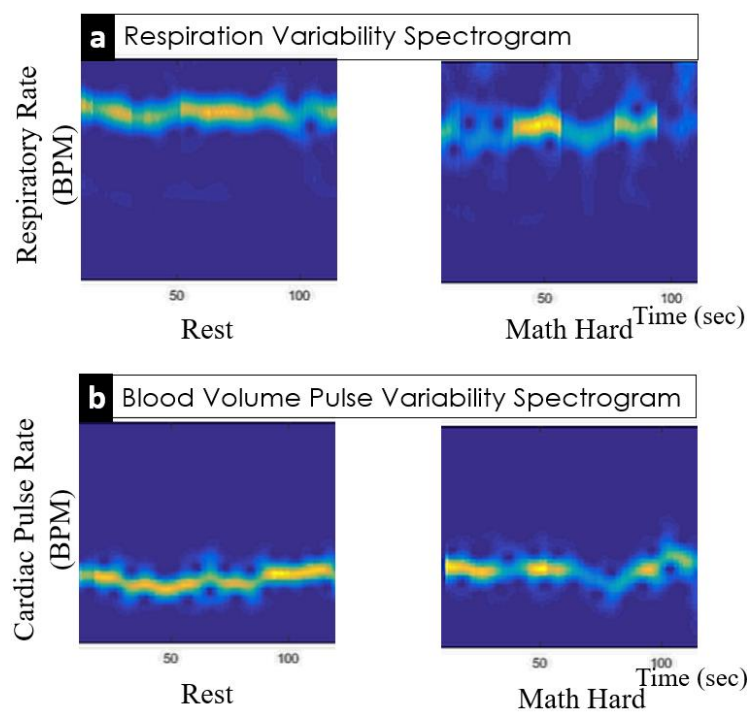


Figure 8.14. Examples of pairs of the respiration variability spectrogram and blood volume pulse variability spectrogram from a rest period and a stress induction task period (the Math hard).

8.8. Chapter Summary

In this Chapter, we have proposed a new system which automatically recognises people's psychological stress level from their breathing patterns which are tracked through the robust respiration tracking method proposed in Chapter 6. This approach contributes in three ways. Firstly, instead of creating hand-crafted features to capture aspects of the breathing patterns, we transform the uni-dimensional breathing signals into two-dimensional respiration variability spectrogram (RVS) sequences. The spectrograms can capture the complexity of the breathing dynamics in a compact form. Secondly, a spatial pattern analysis based on a *Convolutional Neural Network (CNN)* is applied to the spectrogram sequences without the need of hand-crafting features (automatic feature learning). Thirdly, this model has achieved state-of-the-art performance. The proposed method was thoroughly evaluated through a multi-level, multi-task stress induction study - the model was trained and tested with data collected from people exposed to two types of cognitive tasks (Stroop Colour Word Test, Mental Computation test) with sessions of different difficulty levels. Using normalised self-report as ground truth, the proposed system only using breathing information reaches state-of-the-art performance (accuracy: 84.59%) in discriminating between two levels of stress under unconstrained settings (i.e. verbal expressions and movement). As the tracking of the respiration had been thoroughly tested in Chapter 5 for movement and outdoor scenes, we expect this approach will perform as a reliable stress measure in mobile, real-world settings.

Chapter 9

Instant Automated Inference of Mental Stress using Cardiovascular Signatures

In this second part of Stage III, we aim to build another stress recognition system based on vasoconstriction/dilation driven nasal thermal variability which has been explored in Chapter 6. Like the proposed system in Chapter 8 which focuses on respiration variability, we aim to explore how to improve the capturing of nasal thermal variability in association with mental stress. In particular, we explore how this type of cardiovascular thermal signature can contribute to automatic mental stress detection together with other type of cardiovascular signals, HRV. Furthermore, inspired by a finding in the literature that the nose tip thermal variances occur quickly in response to mental load (Abdelrahman et al. 2017), we target a specific task, *instant stress inference*, which only takes a very short time measurement, in turn helping improve real-world use cases of automatic stress detection.

In Section 9.1, we review the background and key challenges. In Sections 9.2 and 9.3, we propose novel recording setup and computational approach for instant stress inference. In Sections 9.4, 9.5, we describe the conducted data collection study, collected dataset and evaluation methods. Section 9.6 provides in-depth analysis of a perceived stress score-based labelling technique which is proposed in Chapter 8. Section 9.7 reports results. This chapter ends with a discussion in Section 9.8 and a chapter summary in Section 9.9.

9.1. Background and Research Question

The literature exploring thermal signatures in relation to mental stress has mostly focused on capturing the binary direction of stress-induced temperature variations in certain facial ROIs.

Amongst such ROIs, the most consistently reported area is the nose tip whose thermal pattern is related to vasoconstriction/dilation (a type of cardiovascular activities) (Genno *et al.*, 1997; Veltman & Vos, 2005; Or & Duffy, 2007; Engert *et al.*, 2014). Whilst it is evident that thermal imaging has the greater potential for automatic stress recognition, the contribution of this type of cardiovascular thermal signature to automatic stress recognition has remained very weak up to date. For example, Engert *et al.* (2014) achieved slightly above chance accuracy in detecting mental stress using thermal directional information from multiple facial ROIs including the nose tip. This result was also limited in terms of highly constrained situations used in the work, which limit one's head movement and control ambient temperature.

Hence, there are needs for a new technique for improving stress detection performance when using such signature and for a method to make this thermal imaging-based approach more feasible in real-world ubiquitous settings. As have been demonstrated throughout this thesis, mobile thermal imaging is a new pathway for physiological measurements. In particular, Chapter 6 has proposed a new method to enable a low-cost thermal camera attached to a smartphone to track nasal vasoconstriction/dilation events. Building on this smartphone-based setup and physiological measurement, we aim to address the discussed issues above.

Another cardiovascular signature, HRV which can be simultaneously measured with vasoconstriction/dilation related nasal thermal signature

Interestingly, recent studies have also demonstrated that it is possible to use smartphone RGB cameras to measure Blood Volume Pulse (BVP) (Jonathan & Leahy, 2010; McManus *et al.*, 2013; Xu *et al.*, 2014; Chan *et al.*, 2016; White & Flaker, 2017). These encouraging results suggest that smartphones could become a powerful apparatus for monitoring and supporting mental stress management on a daily basis through biofeedback (Yu *et al.*, 2018). Indeed, the combination of thermal and RGB cameras into one device has the potential to provide a very large set of physiological measurements for stress monitoring. Given this, we aim to add to this body of work by investigating *two important cardiovascular signals* that can be measured by low cost and low resolution sensors. The main aim is to explore how nasal thermal variability can contribute to stress inference together with another type of cardiovascular patterns.

Heart Rate Variability (HRV) is the time series of variation in heartbeats. It has been used to measure a person's mental stress (Pagani *et al.*, 1986; Camm *et al.*, 1996; Bernardi *et al.*, 2000; Hjortskov *et al.*, 2004; McDuff *et al.*, 2016; Mohan *et al.*, 2016; Jobbágy *et al.*, 2017; Zhu *et al.*,

2017). HRV's popularity arises from the fact that it has been shown to provide information about the sympathovagal balance between the SNS and PSNS. When confronted with a stressor, the autonomic nervous system can produce a sequence of fight-or-flight responses (Everly Jr & Lating, 2012). These manifest themselves as alternations of accelerated and decelerated cardiovascular patterns (Everly Jr & Lating, 2012; Shaffer et al., 2014).

To characterize the HRV, various authors (Pagani et al., 1986; Camm et al., 1996; Hjortskov et al., 2004; Williamon et al., 2013) have proposed a variety of hand-crafted HRV metrics that are computed over the time intervals between heartbeats. Although most of the HRV metrics were originally built based on the R-R intervals from ECG (Electrocardiogram) measurements (Billman, 2011), the metrics have been applied to the P-P intervals from PPG measuring blood volume pulse (Heathers, 2013; McDuff et al., 2016; Mohan et al., 2016; Jobbágy et al., 2017). In the case of PPG, the term Pulse Rate Variability (PRV) or PPG HRV are often used to clarify the different type (even if related) of event measured (Giardino et al., 2002; Schäfer & Vagedes, 2013; Heathers, 2013; Shaffer et al., 2014) with respect to ECG.

Amongst the most commonly used are statistical metrics (such as the standard deviation of R-R or P-P intervals) and frequency-band metrics (e.g. the normalised power in a frequency band of interest). In particular, various studies have found that the Low Frequency (LF; 0.04Hz – 0.15Hz) and High Frequency (HF; 0.15Hz – 0.4Hz) bands of the time intervals in heart rates appear to reflect the SNS and PSNS activities (Camm et al., 1996). Given this observation, many studies have used the LF/HF ratio as a stress indicator (Pagani et al., 1986; Hjortskov et al., 2004; Salahuddin & Kim, 2006; Zhu et al., 2017). However, the single HRV metric as a feature does not strongly contribute to the inference of a person's physiological response to stress (automatic stress recognition using machine learning techniques) as discussed in Chapter 6 and Chapter 8. Hence, multiple HRV metrics have been used together with features from other physiological activities such as perspiration and respiratory activities for automatically inferring mental stress (e.g. for driving tasks in Healey & Picard, 2005 and desk activities in McDuff et al., 2016).

Existing stress recognition systems require relatively long-term physiological measurements, limiting real-world applications

To ensure reliable measurements with such features, a relatively long term window of data (several minutes to a few hours) must be used (McDuff et al., 2016). Given this, the system proposed in Chapter 8 also uses a 2 minutes time window. Although this is acceptable in specialist

settings or with medical devices, it would be highly inconvenient in the real world with unstructured settings using low cost devices. For example, in the case of the smartphone PPG, if a mobile device such as smartphones were to be used, a user would have to continuously make sure their finger is held stably in front of the PPG sensor. Another issue is that changes in light levels, as a user moves around, can corrupt long-term measurements. In addition, they require to measure baseline signals.

The same can be found from studies exploring the relation of the nose tip temperature variations with mental stress levels. To compute the thermal direction, thermal imaging-based studies also requires long-term measurements as well as baseline temperatures, for example, 5 minutes of separate sessions in Engert *et al.* (2014). This limits their use in real-life applications. These limitations led to our research question: *How can the nasal vasoconstriction/dilation related thermal signatures contribute to very fast stress recognition together with other types of cardiovascular signals from another mobile sensing channel (HRV from PPG)?* Accordingly, this chapter focuses on building a novel instant stress inference method which only requires a very short time measurement of both types of cardiovascular signals.

9.2. Recording Set-up

This section presents a method that enables quick inference of a person's perceived stress level with nasal thermal variability (main outcome from Chapter 6) and PPG derived HRV as main inputs to machine learning models. We call these measurements instant measurements to differentiate them from the short measurements (typically between 2min and 5min) which have been previously defined in the literature (Shaffer & Ginsberg, 2017). Below is a description of a recording set-up we implemented to measure nasal thermal variations and PPG-derived HRV profiles simultaneously.

Towards Smartphone as Multiple Cardiovascular Measures

The main cardiovascular sensing channels of this work are the rear RGB camera of a mobile phone (LG Nexus 5) and a low-cost thermal camera (FLIR One 2G) attached to the phone. Figure 9.1 shows the smartphone, the attached thermal camera, the procedure for collecting measurements, and the physiological measurement interface.

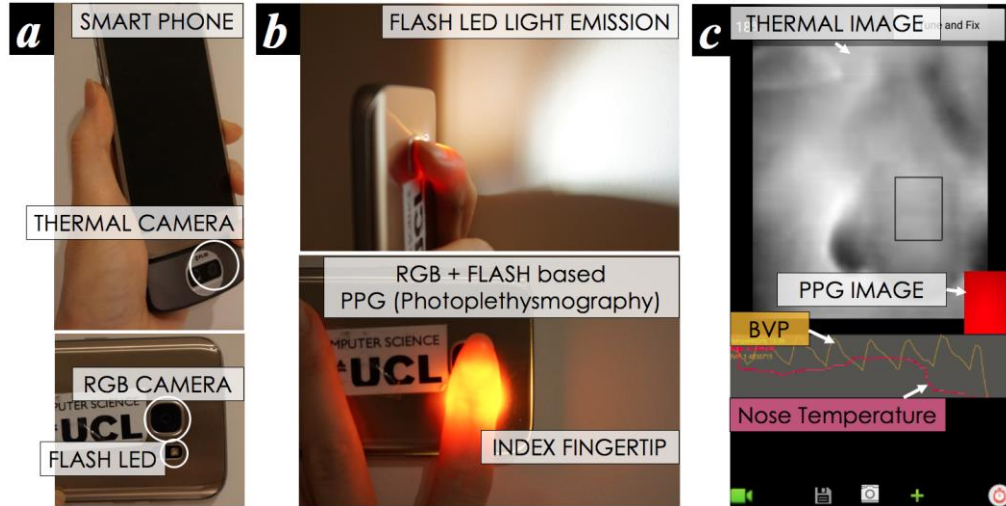


Figure 9.1. Thermal and Smartphone RGB cameras-based physiological measurement: (a) a smartphone with an add-on thermal camera, (b) flash LED emission for PPG measurement, (c) developed software to collect nasal thermal signature from the nose and BVP from the finger.

Although the smartphone-imaging-based PPG measurement can be performed in either a contact (Jonathan & Leahy, 2010; Chan et al., 2016) or a contactless manner (Xu et al., 2014), in our work we only focus on a contact-based imaging PPG. The reason is based upon previously repeated investigation within clinical studies (Chan et al., 2016; White & Flaker, 2017) given its higher accuracy. In addition, given that a normal RGB camera is only sensitive to a narrow electromagnetic spectral range of visible light in the so-called visible spectrum (Cho *et al.*, 2018), adequate lighting is required before it can be used as a PPG sensor. Hence, a light emission from the rear flash LED is used and a user is required to hold the smartphone body and place his/her finger over both the back camera and flash light (Figure 9.1a, 9.1b).

Unfortunately, the use of the back flash limits the duration of the measurements in some devices since its heat can potentially burn a person's skin. As shown in Figure 9.2, a large amount of heat is produced by the LED emission from the chosen smartphone (LG Nexus 5) in just 25-30 seconds of operation. A similar amount of heat was observed from another mobile phone (Samsung Galaxy 6 in Figure 9.1b). Since temperatures above 50 °C are potentially damaging to human skin tissues, for example, skin erythema could occur from 25 seconds heating at 51.07°C (Yarmolenko et al., 2011), we limit the cardiovascular measurement to a 20 second time period. This is also the required minimum duration for obtaining valid HRV metrics values, in particular, LF/HF (Shaffer & Ginsberg, 2017).

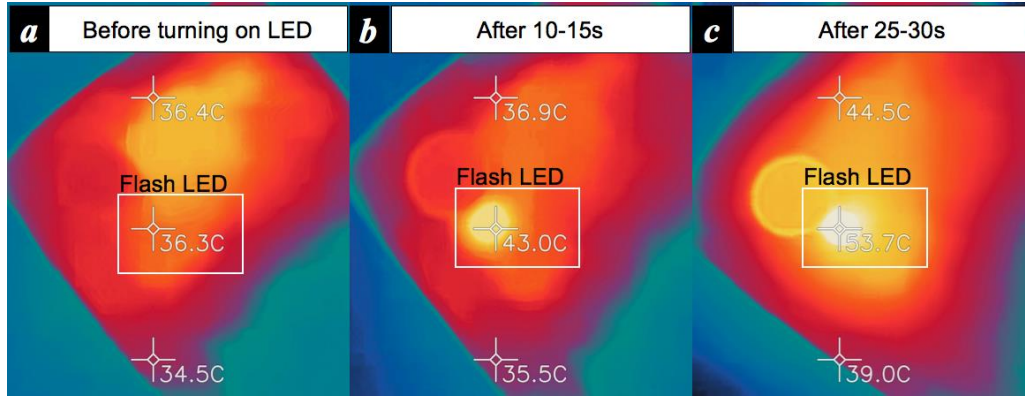


Figure 9.2. Heat produced by the rear flash LED of a smartphone (LG Nexus 5) measured by a thermal camera (FLIR One): (a) before turning on the LED (36.3°C), (b) after 10-15 seconds (43°C), and (c) after 25-30 seconds (53.7°C).

To capture a time series of nasal thermal variability and blood volume pulse sequence, we used the method proposed in Section 6.2 for the former and the method proposed in Appendix F for the latter. The recording interface stores the time stamp with each image frame.

9.3. Proposed Computational Method

In this section, we propose low-level features which can be fed to automatic feature learning frameworks as discussed in Chapter 8.

Low-level Features and Automatic Feature Learning

As in Chapter 8, this approach is for capturing variability of physiological signals as low-level features and having automatic feature learning procedures to learn informative high-level features by themselves. As the 2D representation proposed in Chapter 8 (section 8.2.1) is designed to require relatively long-term measurements, we take a different strategy for signals from very short-term measurements (20 seconds). Figure 9.3 summarises a workflow of low-level features and feature learning. Here we propose to use the *feature-scaled thermal variability sequence* (from the nasal area) proposed in Section 6.2 (see Figure 6.5b) as low-level features. We then use an artificial Neural Network (NN) to automatically learn the mapping between the extracted set of low-level features and stress levels. The use of artificial neural networks can empower automatic learning of informative physiological features with back-propagation to

repeatedly tune internal parameters of a machine learning model to let the features emerge from the data (this is also called representation learning). Likewise, we applied the same strategy to the P-P intervals as low-level features in Appendix F.1 (Figure F.1c), which can be simultaneously produced by the developed interface (Figure 9.1c). This low-level feature set can represent variability of cardiovascular activities.

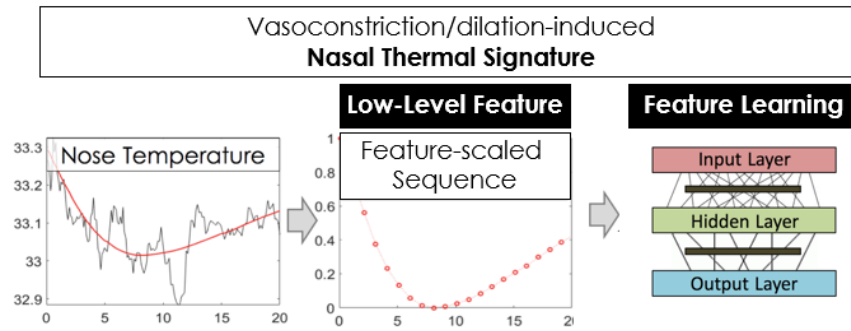


Figure 9.3. Low-level features and feature learning: we take the feature-scaled thermal variability sequence proposed in Chapter 6 as low-level features which are fed to an automatic feature learning model using a single-hidden layer neural network.

9.4. Experimental Protocol and Dataset (*InstantStress*)

Based on the experimental protocol used in Chapter 8, this study protocol consisted of:

- waiting in the corridor, introduction and entering the study room (5-10 min)
- information/consent/demographics forms filled in (5-10 min)
- Session 1
 - [Rest 1] **sitting, resting** (5 min)
 - 20s measurement and self-reporting of perceived stress (1-2 min)
 - [Task 1] **Stroop Test 1** (5 min)
 - 20s measurement and self-reporting of perceived stress (1-2 min)
 - break (5 min)
 - [Task 2] **Stroop Test 2** (5 min)
 - 20s measurement and self-reporting of perceived stress (1-2 min)
 - break (3 min)
- Session 2
 - [Rest 2] **sitting, resting** (5 min)
 - 20s measurement and self-reporting of perceived stress (1-2 min)
 - [Task 3] **Math Test 1** (5 min)
 - 20s measurement and self-reporting of perceived stress (1-2 min)
 - break (5 min)
 - [Task 4] **Math Test 2** (5 min)
 - 20s measurement and self-reporting of perceived stress (1-2 min)
 - break (5 min)
- Closing
 - wrap-up and participant's feedback (5-20 min)

As described above, heat caused by the use of the smartphone PPG limited our data gathering to a 20 second window immediately after each task. The aim is to capture the cardiovascular changes related to stress responses and their dynamics immediately after the stressor has ended instead of measuring the signals during each task (Figure 9.4). For the 20s physiological measurements, the person was asked to hold their index finger on the smartphone RGB camera while keeping the smartphone add-on thermal camera facing their nose, as shown in Figure 9.4b. We call the dataset collected from this instant measurement, *InstantStress*.



Figure 9.4. Experimental setup and self-report question: (a) during each stress-induction task session, (b) 20 second physiology measurement after sessions.

A total of 17 healthy adults (mean age 29.82 years, SD=12.02; 9 female) of varying ethnicities and different skin tones (pale white to black) were recruited from the University College London and non-research community through the UCL psychology subject pool system. Participants completed pre-screening through the system which was designed to exclude participants with any history of psychiatric disorders or medicine intakes which may influence their physiological signatures. The study was conducted in a quiet lab room with no distractions and without controlling the room temperature.

9.5. Evaluation Method

First, to understand how existing metrics of cardiovascular signatures work properly in instant stress measurements, we tested correlations between self-reported stress ratings (which we collected using VAS-based questionnaires in Appendix H) and metrics values. As for metrics for nasal thermal signatures, together with the most widely used metric to capture thermal directionality (Temperature Difference, TD), we also used two further metrics which we proposed in Section 6.2: Standard Deviation of Successive differences of Thermal Variability sequence (SDSTV) and Standard Deviation of Thermal Variability sequence (SDTV). We used feature-scaling-based metrics as they performed better than original metrics in Chapter 6 (i.e. TD_F , $SDSTV_F$, $SDTV_F$ in Table 6.1). As for metrics for blood volume pulse from PPG, we followed earlier studies on stress inference using HRV metrics (in our case, PPG HRV or PRV metrics) (Healey & Picard, 2005; Wang et al., 2013; Hovsepian et al., 2015; McDuff et al., 2016): LF Power, HF Power, LF/HF ratio, SDPP (Standard Deviation of P-P intervals), RMSSD (Root Mean Square of the Successive Differences of P-P intervals) and pPP50 (proportion of the number of the successive differences of P-P intervals greater than 50ms *of the total number of the intervals*).

Second, to evaluate stress inference performance of our proposed approach, we used such metrics as *high-level features*:

- 1) *Thermal F1*: TD_F , *Thermal F2*: $SDSTV_F$, *Thermal F3*: $SDTV_F$,
- 2) *HRV F1*: LF Power, *HRV F2*: HF Power, *HRV F3*: LF/HF ratio, *HRV F4*: SDPP, *HRV F5*: RMSSD, *HRV F6*: pPP50.

With these features, we used the k-Nearest Neighbour algorithm (denoted as kNN, $k=1$) as a benchmark stress inference model given that this is typically used in this area (Wang et al., 2013). By choosing this, we aim to assess the limitations of the use of handcrafted features which may simplify a person's dynamic physiological events, and in turn possibly miss out some fast, informative moments. In particular, in the case of instant measurements (short period of time), this cannot be compensated by the use of a sliding window producing sequential feature values (e.g. a 120 seconds sliding window used in McDuff et al., 2016 to continuously produce PPG-HRV features during a 180 second task session).

For the implementation of the neural network, we tested two sizes of hidden layer nodes: a) small ($n=80$, NN1) and b) large ($n=260$, NN2) – each node size was empirically chosen. The mean and standard deviation of the training dataset were used to normalise both the training and

testing dataset. The sigmoid was used as an activation function. In the training process, a fixed learning rate of 0.5 was used for 100 epochs.

9.6. Perceived Mental Stress Levels and Labelling

Given the focus on automated inference of a person's perceived stress level, the labelling of self-reported stress scores is an important step. However, as discussed in Chapter 8 interpersonal variability has been repeatedly found from self-reports of perceived mental stress (Hernandez et al., 2011; Zhu et al., 2017). This is a key issue which must be addressed if we are to create automatic stress recognition systems that can generalise across people.

Using the labelling technique proposed in Chapter 8, we normalise all perceived stress scores collected from each participant through feature scaling that identifies the minimum and maximum scores for a participant and rescales all the scores so the range is the same across all participants. Then, the K-means algorithm ($k=3$) is used to group the participants' VAS scores into three levels of perceived stress scores corresponding to "None or low stress", "Moderate" and "Very high" on the VAS we used (see Figure 3.4 in Section 3.3). In this chapter, we focus on discriminating between two levels of stress, *No-stress* and *Stress* given the limited amount of data for a more refined discrimination. Hence, a third step is required. We split the labels into two groups: the *No-Stress* group referring to the K-mean "None or low stress scores" cluster and the *Stress* group containing both the K-mean "Moderate" and "Very high" score clusters. Two obtained labelled groups are hence used to label the related physiological signatures from each 20s window (L1).

Furthermore, we explore the possible effect of different data labelling strategies: a) L2: combining the first and second K-means clusters (from $k=3$) into No-stress by contrast with L1, b) L3: K-means with $k=2$, and lastly, c) L4: the original stress scores divided by directly dividing the VAS scale into three equal sections and then combining the "moderate" and "Very high" stress classes into one, i.e. "Not at all" and "Moderate + Very high" (threshold at point 3.334 on the VAS scale in Figure 3.4). The aim of L2 and L3 was to understand the sensitivity of our approach in separating the moderate level of stress with the other two classes. L4 was used as a way to compare with more standard techniques used in the field (Sano & Picard, 2013).

9.7. Results

From the 17 participants, we collected 102 sets of thermal variability sequences, estimated BVP and P-P intervals from *20s-instant measurements* taken after each Stroop and Math task and after each resting session. However, 2 sets of data were not recorded due to phone battery issues at the end of one experiment, and 1 set was not recorded as one participant clicked the turn-off button on the phone by mistake. 6 further sets had to be discarded because some participants' nose was not visible by the thermal camera (nose outside of the range of view due to sudden severe coughing during the 20sec, or because of head turned towards the experimenter, or the nose was covered by a person's hand). Although these disturbances were often transient, they meant that data could not be collected within the 20s immediately following the end of the stressor.

An analysis of the thermal data from Rest 1 also showed some extreme patterns in the nose tip temperature (e.g. sudden increase in temperature). This may be explained by the fact that the experiment was conducted during the winter and temperatures outside of the experimental room were often significantly lower. This included both outdoors, and indoors in the corridor where the participants waited for the experiment. Despite the temperature changes, the Rest 1 data was kept in the dataset. A total of 93 sets were used for the study.

Self-reported stress ratings and hand-engineered metrics

An important step was the analysis and possible normalisation of the self-reported stress scores. The boxplot in Figure 9.5 shows the distribution of the self-reported scores over the resting periods and the different sessions and tasks. Overall, the stress elicitation procedures produced the wanted levels of stress with the hard sessions scoring higher than the easy sessions and the latter scoring higher than the resting periods (Rest from Session 1: $M=1.49$, $SD=1.94$; Rest from Session 2: $M=1.30$, $SD=1.26$; Stroop Easy: $M=2.17$, $SD=1.46$; Math Easy: $M=2.66$, $SD=1.80$; Stroop Hard: $M=3.92$, $SD=2.11$; Math Hard: $M=5.17$, $SD=2.55$) despite two outliers. However, the wide boxplots also show inter-subject variability in self-reporting. In addition, the ranges (maximum - minimum) in scores for each participant differ quite highly (Maximum range: 8.75, Minimum range: 1.5, Mean: 4.7, Std: 2.1) further suggesting the need for normalisation of the scores.

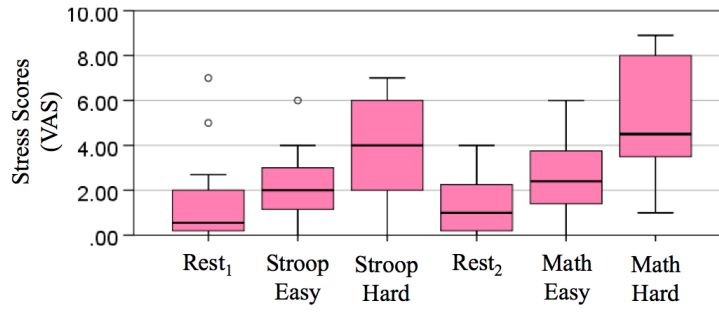


Figure 9.5. (Boxplot) Inter-subject variability shown from the original self-reported stress scores of the 17 participants (box plot, 95% confidence interval) across each section (Rest₁, Stroop Easy, Stroop Hard, Rest₂, Math Easy, Math Hard).

Therefore, the data for each participant was normalised with respect to their range of scores over all the sessions. Figure 9.6a shows the original data and Figure 9.6b shows the normalised data. The normalisation helps to identify two main modes in the score distributions suggesting the presence of two main clusters of stress levels. Given the subjectivity of stress ratings and the limited amount of data sets to carry a multi-level model, in this chapter we focused on binary classification of perceived mental stress: no/low stress vs medium/high (or very high) stress. To this end, we used the K-means based labelling technique to cluster the normalised perceived scores into two groups as described in Section 9.6. The K-means separation between the two clusters is represented by each different colour in Figure 9.6b

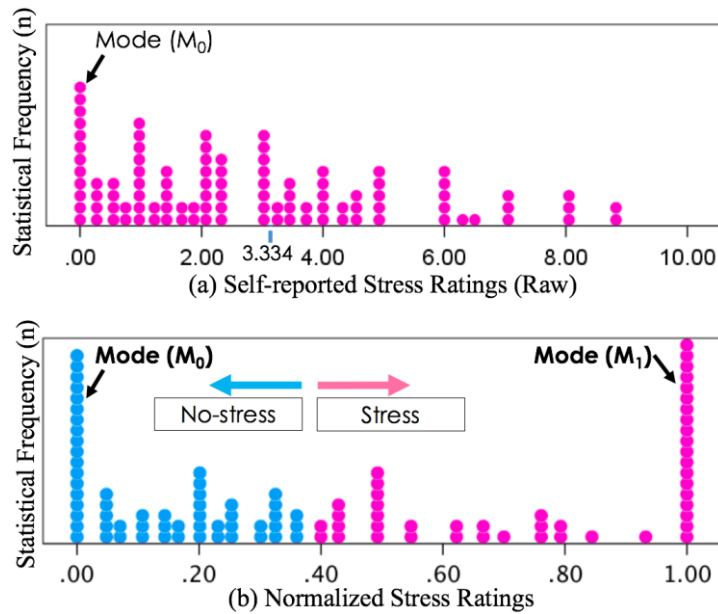


Figure 9.6. Overall self-reported stress score distributions (from 17 participants over the sessions including the resting periods) - (a) original scores, (b) normalised stress scores (normalisation of scores from each participant) clustered into No-stress and Stress groups along with outputs of K-means.

We tested the correlations between the original self-reported scores, normalised self-reported scores and the high-level features extracted from the HRV and thermal signatures as summarised in Table 9.1 (using Pearson correlation coefficients). The normalised self-scores maintained a high correlation with the original scores ($r=0.752$, $p<0.001$). While some features from each physiological sensing channel were significantly correlated between themselves (e.g. HRV F2 - F4: $r=0.838$, $p<0.001$; Thermal F1 - F3: $r=0.803$, $p<0.001$), the correlation values were lower across sensing channels. In addition, only Thermal F2 features show approaching significance but low correlation with the self-report scores ($r = .196$, $p=0.059$), indicating that each individual engineered feature alone could not lead to high discrimination between perceived levels of stress.

Table 9.1. Pearson correlation coefficients across self-reports, PPG derived HRV and thermal high-level features. S1=Normalised self-reported scores, S2=original self-reported scores.

			Self reports		PPG derived HRV						Thermal		
			S1	S2	LF (F1)	HF (F2)	LF/HF (F3)	SDPP (F4)	RMSSD (F5)	pPP50 (F6)	TD (F1)	SDST V (F2)	SDTV (F3)
Self Rep.	S1	Corr.	1	.752	0.007	0.011	-0.044	0.03	0.146	0.058	-0.154	0.196	0.02
		p-value		<0.001	0.943	0.911	0.665	0.77	0.148	0.569	0.139	0.059	0.848
	S2	Corr.		1	-0.079	-0.044	-0.082	-0.002	0.083	0.097	-0.153	0.197	0.032
		p-value			0.438	0.664	0.422	0.987	0.414	0.338	0.14	0.058	0.758
PPG derived HRV	F1	Corr.			1	0.394	0.573	0.638	0.098	0.134	0.016	0.12	0.047
		p-value				<0.001	<0.001	<0.001	0.336	0.186	0.88	0.251	0.657
	F2	Corr.				1	-0.293	0.838	0.13	0.39	0.083	0.2	0.054
		p-value					0.003	<0.001	0.199	<0.001	0.431	0.054	0.608
	F3	Corr.					1	0.007	-0.027	-0.178	0.056	0.057	0.123
		p-value						0.948	0.791	0.079	0.596	0.588	0.239
	F4	Corr.						1	0.139	0.571	0.1	0.198	0.084
		p-value							0.171	<0.001	0.338	0.058	0.425
	F5	Corr.							1	-0.067	-0.059	0.174	-0.067
		p-value								0.511	0.572	0.095	0.521
	F6	Corr.								1	0.134	0.212	0.127
		p-value									0.2	0.042	0.225
Thermal	F1	Corr.									1	0.213	0.803
		p-value										0.039	<0.001
	F2	Corr.										1	0.487
		p-value											<0.001
	F3	Corr.											1
		p-value											

Figure 9.7 shows values of each pre-crafted metric across the sessions (rest and four stressful events, i.e. Stroop easy/hard and Math easy/hard) and across the labels produced by the labelling technique. As shown in Figure 9.7a, there was no common pattern found between two easy or hard tasks, although they were designed to induce similar levels of mental stress (e.g.

easy: low stress level, hard: high stress level). For example, Thermal F1 appeared to strongly decrease during the Math hard task but not during the Stroop hard task, Thermal F2 increased with the Stroop hard task, but less during the Math hard task. HRV F5 was generally high after both Math easy and hard task sessions than Stroop hard session. This can indicate further that each feature alone from the instant measurement is less likely to contribute to the inference of each session. On the other hand, when we applied our labelling technique, Thermal F1 values grouped into stress were generally lower than no-stress data as shown in Figure 9.7b, which were consistent with findings from literature (Genno et al., 1997; Or & Duffy, 2007; Engert et al., 2014).

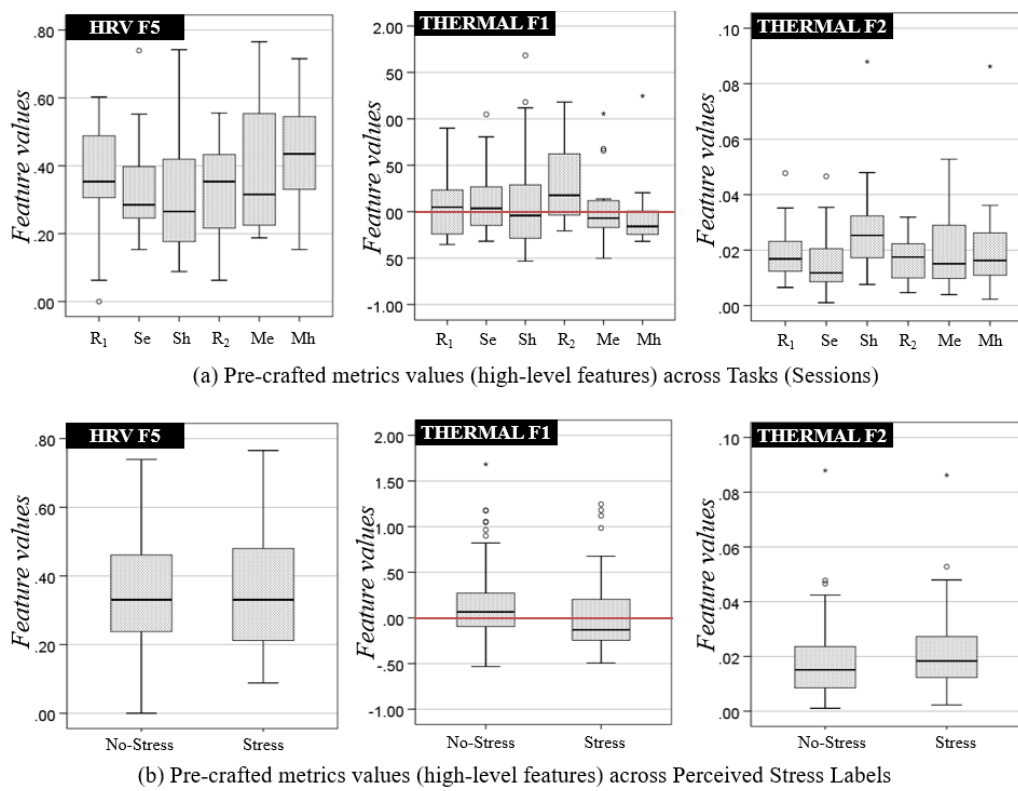
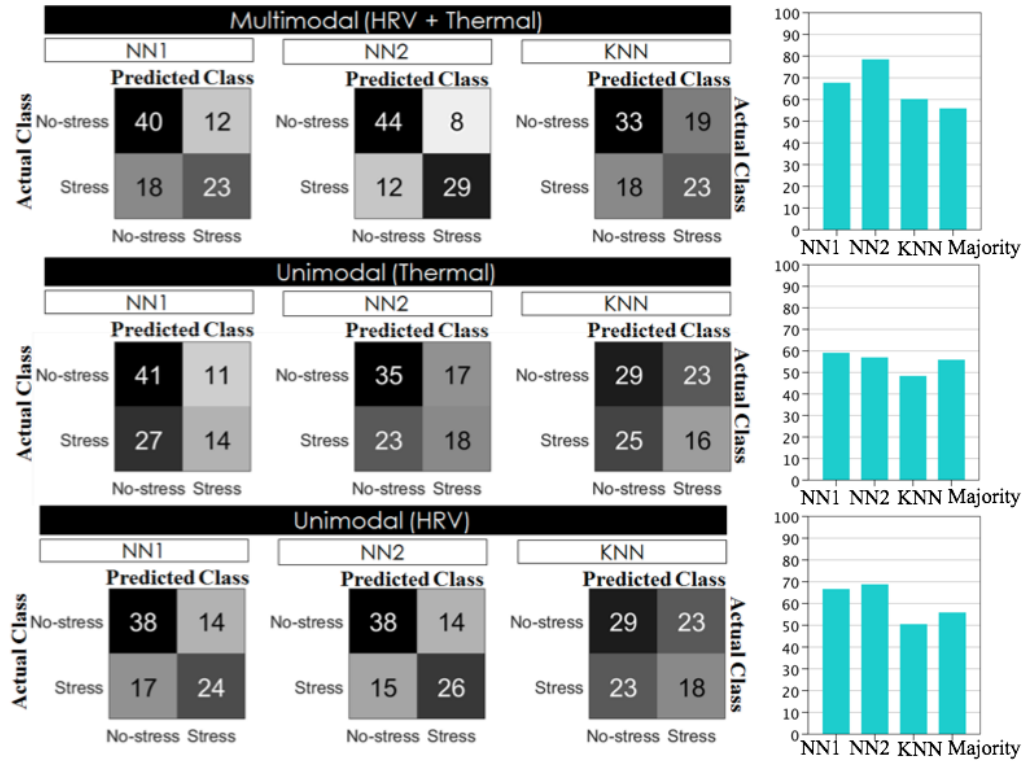


Figure 9.7. Box plots of 95% confidence intervals in values of each pre-crafted metric across (a) each session (R₁: Rest 1, Se: Stroop easy, Sh: Stroop hard, R₂: Rest 2, Me: Math easy, Mh: Math hard) and (b) label produced by our labelling technique. The three features (having best correlations with self-reports) are HRV F5: *RMSSD*, root mean square of the successive differences of P-P intervals, Thermal F1: *TD*, temperature difference between from the start and the end (a red line is drawn to show negative or positive thermal direction), F2: *SDTV*, standard deviation of the successive differences of thermal variability sequence.

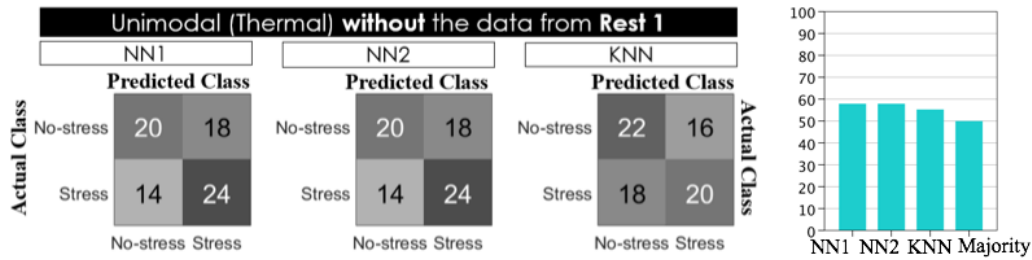
Instant Stress Inference Results

To evaluate the performance of instant stress recognition, we used a 17-fold leave-one-subject (participant)-out (LOSO) cross-validation (see Figure 8.6). LOSO was chosen to test the ability to generalize to unseen participants (*one size fits all*) as in Chapter 8. Figure 9.8 and 9.9 summarise the accuracy results of the three classifiers (NN1, NN2, kNN) using LOSO (N=17) for three different cases: a) multimodal-approach by simply combining features from both sensing channels (HRV, Thermal), b) unimodal approach using thermal features, and c) unimodal approach using HRV features. Both neural networks NN1 and NN2 used our proposed low-level features only (i.e. P-P intervals and thermal variability sequences). Data sets were not balanced (No-stress: 52, Stress: 41 instances), leading to the ratio of the number of data in the majority class to the total number of dataset (52/93=55.91%). Overall, the NN2-based multimodal approach produced the highest mean accuracy of 78.33% (SD=15.43) (mean F1 score=77.92%) in discriminating between no-stress and stress (see confusion matrix in Figure 9.8a for details). The NN1 (whose hidden layer is smaller than that for NN2) produced a lower accuracy (M=66.76%, SD=21.75). From all cases of modality, the kNN with the high-level features (i.e. the hand-engineered 6 HRV and 3 thermal features) performed worst. A similar pattern can be seen for the HRV unimodal channel (NN1: M=65.78%, SD=20.55; NN2: M=68.53%, SD=18.89; kNN: M=50.20%, SD=19.63). For the thermal channel, the NN1 appears to perform marginally better (M=58.82%, SD=21.11) than the NN2 (M=56.67%, SD=18.79), but again both NNs performed better than the kNN (M=48.14%, SD=16.52).

However, it should be noted that, for all the models, the confusion matrices for the thermal case (Figure 9.8a - Thermal) show a clear bias towards the no-stress class. Given this bias and the fact that thermal data from the Rest 1 sessions appeared to be affected by the large variation in temperature between the waiting space and the experiment room (in addition, some participants had just arrived from the outside while others had been already indoor for sometimes), we re-ran the models discarding the data from the Rest 1 sessions. Whilst the overall performance over this modality did not change largely (NN1: M=58.14%, SD=23.33; NN2: M=58.14%, SD=21.59; kNN: M=55.88%, SD=22.38) and the NN1 and NN2 still performed better than the kNN with hand-engineered features, all the confusion matrices (Figure 9.8b) show more balanced results and a better prediction of the stress class overall.



(a) Confusion Matrix and Accuracy (for Multimodal, Unimodal cases)



(b) Confusion Matrix and Accuracy (for Thermal without Rest1)

Figure 9.8. Summary of accumulated (from 17 fold testing sets - LOSO) confusion matrices and accuracies (a) for the three classifiers NN1, NN2 and kNN along with each set of modalities (Multimodal: HRV+Thermal, Unimodal: Thermal, HRV), (b) for the temperature-based unimodal approach built without the Rest 1 data. Each number in the confusion matrices refers to the number of instances.

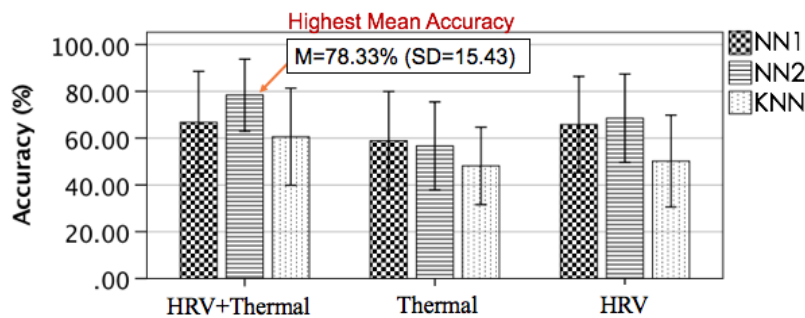


Figure 9.9. Summary of mean inference accuracy results (testing) across 17 folds for the three classifiers NN1, NN2 and kNN along with each set of modalities (Multimodal: HRV+Thermal, Unimodal: Thermal, HRV)

Lastly, we investigated the effect of the normalisation and K-means clustering of self-reported scores in inferring the perceived stress levels. For this part of the study we removed the Rest1 data. There were two reasons for this. First, we wanted to avoid the noise from those set of data affected the comparison between the labelling methods. Second, this was also to obtain a more balanced number of instances in each class for testing different labelling methods, less biasing the learning process. The comparison of models over the different labelling techniques did not aim to obtain better performance, but to understand how normalisation and different clustering approaches could affect the modelling by acting on class separation and inter-person variability in subjective self-reports. We were also interested in understanding how sensitive the system was in separating stress scores by using the same dataset and merging the intermediate levels with one of the two classes (L1 and L2).

We tested the three models (NN1, NN2, kNN) for the multimodal-approach with the different labelling strategies (L2-L4, introduced in the previous section). Figure 9.10 summarises the accuracy results for four different strategies - L1: the main method, L2: K-means with $k=3$, but combining no-stress and moderate level stress scores as one group, L3: K-means with $k=2$, dissecting the moderate level scores into no-stress and stress, and L4: original scores divided by a point between no-stress and moderate levels (i.e. 3.334 of 10, see Figure 3.4 in Section 3.3). The results showed that the L1 performed best in separating the bimodal distribution of normalised self-reported scores and helped address the inter-personal variability issue. Indeed, all three models obtained the best accuracy with L1 and the worst performance for L3 and L4 with L4 being marginally better than L3. Finally, it should be noted that in the case of L3 and L4, the best performance was obtained with NN2 rather than NN1. This may indicate that mapping feature values to perceived stress scores may benefit from a larger hidden layer to capture the complexity of the relation.

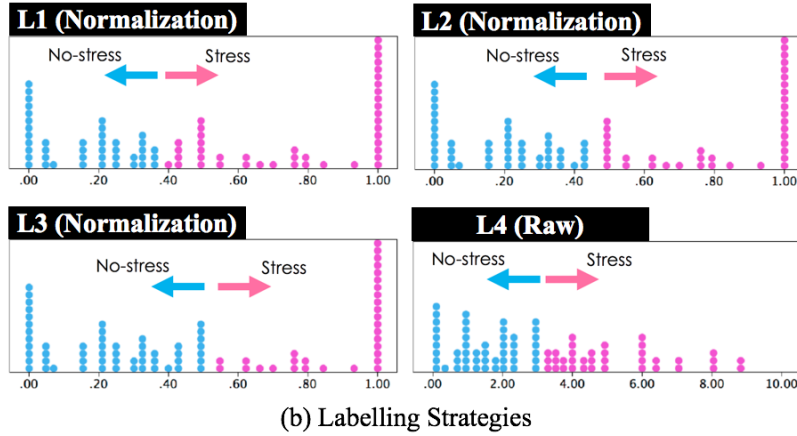
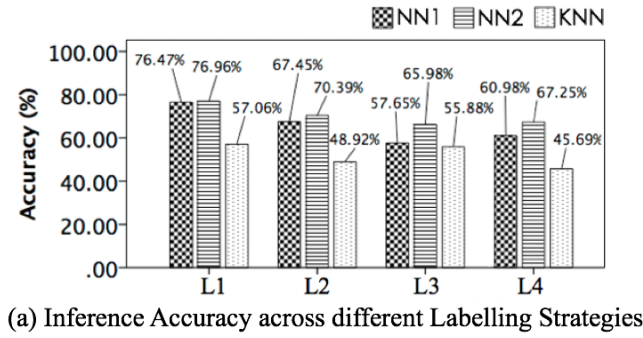


Figure 9.10. Summary of (a) inference accuracy along with (b) different labelling approaches (L1: K-means with $k=3$ and combining moderate and high stress scores, L2: K-means with $k=3$ and combining no-stress and moderate level stress scores, L3: K-means with $k=2$, L4: original scores divided by the border between no-stress and moderate levels).

9.8. Discussion

In this chapter we have focused on how to quickly capture variability of vasoconstriction/dilation related cardiovascular signature in order to build a novel instant stress inference system. In addition, we have explored how this strategy can improve performance of stress detection models using another sensing channel when combined with other types of cardiovascular signature, PPG-derived HRV. Accordingly, we have mainly contributed (i) an exploration of the role of existing metrics in instant inference tasks, (ii) a novel instant stress detection model based on a novel feature set to improve the capturing of nasal thermal variability and heart rate variability, (iii) a further investigation of effects of clustering self-report scores to take into account the subjectivity of the self-report and ensure clear separation between the level of stress to be modelled. Below are detailed discussions of these contributions.

Traditional cardiovascular metrics do not capture stress-related variability from an instant measurement

The contribution of the widely used thermal metric (i.e. TD: Temperature Difference) from the short-term measurement to stress quantification was very weak. This may explain inconsistent findings in the literature where the single metric has been used to capture thermal responses to stressful events (Veltman & Vos, 2005 and Engert *et al.*, 2014). Capturing dynamical information of skin temperature may help improve our understanding of its complex phenomenon.

Similarly, we found that the capability of the hand-engineered HRV metrics, which were used as high-level features in the literature (McDuff *et al.*, 2016; Mohan *et al.*, 2016; Jobbágy *et al.*, 2017), in instantly quantifying stress was very limited (see Table 9.1). It should be noted that we used PPG-derived metrics rather than the more investigated ECG-derived metrics in this chapter. However, strong correlations have been found between the two signal metrics in the case of healthy participants and limited physical movement (Heathers, 2013; Plews *et al.*, 2017). Furthermore, it should be noted that stressors in general affect cardiac pulse-related events even if the two types of events (heart rate and blood volume pulse) may be differently affected within non-healthy or elderly population and extreme situations (hot temperature) (McKinley *et al.*, 2003; Allen, 2007; Shin, 2016).

New representations of nasal thermal variability and heartrate variability contribute state-of-the-art performance in instant stress inference

On the basis of the low correlation between perceived mental stress levels and typically engineered metrics for these two signals, we proposed to use thermal variability and P-P interval sequences as a novel set of low-level features to capture stress responses of cardiovascular activities. With this, we investigated how to benefit from automatic feature learning capabilities of machine learning classifiers (i.e. NNs) in instantly inferencing mental stress. The results showed clear improvements in performance. Indeed, our proposed method with the two cardiovascular signatures achieved 78.33% correct recognition accuracy (NN2 model in Figure 9.9), whereas the kNN with the widely used high level features produced much lower accuracy (60.59%). Similarly, using the HRV-related features only, there was an improvement by 18.33% with respect to the traditional approach (50.20%). The improvement on the thermal channel was smaller but still evident from the results.

In addition, two further contributions can be highlighted from our approach to the modelling of automatic stress inference: *instant measurements* and *no need for baseline*. First of all, previous work required relatively long-term measurements of between 2 to 5 minutes (Tsuji et al., 1996; Engert et al., 2014; McDuff et al., 2016). Indeed, our results demonstrated the possibility to use just a 20 seconds measurement to automatically discriminate between stress and non-stress moments. This approach achieved state-of-the-art performance when compared with approaches using much longer measurements (up to around 70 - 80% correct recognition from LOSO cross-validation; e.g. Hernandez et al., 2011). This is very important given that stillness is critical during PPG measurements and to a certain extent for thermal imaging. In fact, even if automatic ROI tracking methods may help with thermal measurements, people tend to easily move away from the camera or cover their nose with their hands (5 participants did so at least once even for 20 seconds).

Lastly, our approach (more reliable signal and richer features) led to state-of-the-art results without the use of a baseline. This is critical to everyday life settings as in everyday life such baselines may be difficult to establish. Resting periods just before a stressful event cannot be planned and continuously gathering such measures can be costly, whilst at the same time non-stressful resting periods would also need to be automatically detected. In addition, our data from resting periods shows that such a gold standard resting situation does not exist and environment temperature may change drastically, affecting skin temperature. This could have been due to a lab effect, but general everyday situations may also have specific effects on the data. In the existing works, even when using differential features (e.g. temperature differences between two areas of the face - forehead and nose tip), a baseline period was used (Abdelrahman et al., 2017). The lack of baseline is overcome here by proposing richer features capturing informative physiological variations over time.

How do we define the ground truth: what is the best approach?

Setting the ground truth is a difficult process when dealing with subjective reports. How to use self-reports to label the data is a critical issue in the field due to their subjectivity. Interpersonal variability has been repeatedly reported as a critical barrier for building stress inference or quantification systems that can generalize across people (Hernandez et al., 2011; Zhu et al., 2017). The inter-subjectivity of self-reports and the need to reduce the number of classes along with types of applications or the size of the dataset require some decisions on how to refine the labels to be taken. In doing so, there is the danger to add noise to the dataset and

hence to the modelling process. We explored how different labelling techniques may affect the modelling process.

We proposed to address this problem. The first step was to use a standard normalization technique to take into account personal score ranges over all tasks that aimed to induce a wide range of stress levels (from none to medium to quite high). This transformation led to a bimodal distribution highlighting at least two opposite levels of stress (low and high - see Figure 9.6b), whilst it still maintained their strong correlation with the original scores ($r=.752$, $p<0.001$). The bimodal distribution is interesting as, given the low number of participants, it suggests the moderate level of stress is not well separated from the other two classes. A binary classification was hence a sensible approach to take in this chapter, however with larger datasets a more refined analysis and modelling should be carried out. Second, we used a machine learning clustering technique, K-means, to improve separation of the scores into two classes of stress. The results obtained from the comparison of our approach (L1) with its variation (L2) and the more typically used approaches (L3 and L4) led to an interesting lesson on how to create a more reliable ground truth rather than increase noise in labelling.

Then, how should the data be clustered?: according to the number of stress levels to be recognized, or according to the number of stress levels the data collection experiment was set to induce? The latter approach appeared to be more successful. All labelling methods using $K=3$ (L1, L2 and to a certain extent L4) performed better than L3 using $K=2$. This suggests that directly clustering according to the number of classes to be recognized (2 in our case) may spread instances with similar stress level responses (in this case medium responses) across classes introducing noise rather than overcoming the problems of intersubjectivity. However, it should be noted that the normalization step was important. Indeed, the models built on either L1 and L2 using the normalized scores performed better than L4 where the original scores were used instead.

Another important issue to be addressed is: how should the data be grouped when the number of classes to be detected is smaller than the number of levels induced? This decision could be needed either because there were no sufficient instances for a more refined inference or because the application at hand did not require such level of granularity (at the risk of introducing noise due to intersubjective variability). The results showed that L1, collapsing the moderate level with the high one into one class, led to better performance than L2 where medium and no/low stress scores were instead combined. This may suggest that, unless the stress level is very low, stress responses share more similarities than with no-stress responses. A more in-depth analysis

of this aspects could be part of a future work and it may require an in-depth analysis of individual responses and validations over other datasets.

Whilst the results provide some interesting insights on how to cluster data from experiments, a question remains on how to deal with data from real-life situations. It is expected that in real-life situations larger datasets may enable finer levels of discrimination personalized to a specific person. In such situations, as the dataset grows, parameters for labelling may need to be adapted to optimize the personalization. However, such rules we used could be helpful to bootstrap models on the basis of experimental datasets or well-structured initial real-life data collections. The bootstrapped models could then be personalized to specific users and recognition levels as data would be continuously collected by the person.

Limitations and Future Directions

Despite the findings and contributions described above, there is still space for improvement. First, our proposed approach did not perform properly on multiple levels of stress (labelling the data using perceived self-scores). As discussed, this was most probably due to the limited size of the dataset, especially for the medium level of stress (out of three levels). Deploying built software in real life could be a way to build a larger dataset (Mehrotra *et al.*, 2017). With a function to collect self-reported person's perceived stress scores (e.g. digitalized VAS sliding bar in an app), this data collection in the wild could produce a sufficient size of cardiovascular signals sets to support more reliable performance in inferencing multiple levels. In addition, it would be interesting to investigate how the transformation of the self-reported scores could be used to support multi-class classification.

Secondly, this work focused on sedentary situations (but without constraining one's mobility) and did not include physical activity (e.g. walking). It is well known that physical activity induces cardiovascular changes, in turn affecting stress inference performance (e.g. Hong *et al.*, 2012). Hence, it would be interesting to test the instant stress inference ability of our system in situations where there is a considerable amount of physical activities, for example, industrial factory work floor.

9.9. Chapter Summary

In this chapter, we have explored the contribution of nasal thermal variability to automatic stress recognition together with other type of cardiovascular measure, heart rate variability. In particular, with the long-term aim of building a stress monitoring system for mobile, everyday use, the approach has focused on how to quickly infer a person's perceived level of stress from instant physiological measurements.

Building on the findings from Chapter 6, we have investigated (i) the performance of high-level features from nasal temperature in instant inference tasks, (ii) then proposed novel low-level features to represent thermal variability, and (iii) built perceived stress recognisers based on unimodal and multimodal channels. Through the data collection study with 17 participants and a series of stress inducing tasks with different levels, we have demonstrated how this approach is able to achieve state-of-the-art performance using 20 seconds of data, rather than the 2 to 5 minutes required by existing methods. Furthermore, we have investigated effects of clustering self-report scores (to build the ground truth) to take into account the subjectivity of the self-report and ensure clear separation between the level of stress to be modelled.

Chapter 10

Conclusions

The aim of this thesis was to propose a new pathway to thermal imaging-based physiological and affective computing, addressing limitations of *static thermal imaging in highly constrained settings* - which has been dominantly used in this area - and in turn providing practical solutions that work in HCI in-situ and in mobile contexts. Towards this goal, we have explored how to reliably extract multiple physiological thermal signatures and to automatically monitor a person's mental stress levels in less constrained, more real-world like settings.

Built on the diagram summarising findings from the literature, which we have elaborated in Chapter 2 (Figure 2.2), Figure 10.1 illustrates our contributions we have achieved throughout the proposed three-stages approach. Below is a summary of the key contributions filling the gaps that have emerged from the literature.

STAGE I. Novel Signal Processing Techniques for Mobile Thermal Imaging

In Chapter 4 we have identified the effect of environmental temperature which is one of the key issues keeping thermal imaging from being used in unconstrained, mobile settings. To address this, we have proposed a novel *pre-processing technique that adaptively converts temperatures into thermal images* so that graphical components can remain more consistent against variable thermal scenes, in turn improving automated ROI tracking performance.

STAGE II. Robust Physiological Computing through Mobile Thermography

In Chapters 5 and 6, we have proposed novel computational methods capable of automatically tracking *respiratory* (Chapter 5) and *vasoconstriction/dilation related cardiovascular* (Chapter 6) thermal signatures and computing their rich metrics more reliably in

unconstrained settings. In Chapter 7, we have explored novel *cortical thermal signatures*, from underexplored scalp local areas. Accordingly, we have proposed methods for cortical mapping, timeseries data extraction, and metrics for capturing cortical thermal variability, which can foster this new brain study tool within this community.

STAGE III. Improved Automatic Recognition of a Person's Mental Stress based on Mobile Thermography

In Chapter 8, we have proposed a novel representation of respiration that helps condense complex stress-related respiration variability with two-dimensional data, from which automatic feature learning (e.g. deep CNN) methods can search for informative features. The built respiration-based stress recognition system has achieved state-of-the-art performance despite the use of only the breathing signature. In Chapter 9, we have proposed a novel instant stress detection system which combines vasoconstriction/dilation data with another type of cardiovascular signature, PPG-derived HRV. By focusing on capturing physiological variability, the system has also achieved very good performance for inferring mental stress levels.

Across the three stages, this thesis also contributes to datasets of thermal data in both constrained settings and everyday settings. The datasets were incrementally built according to each type of analysis conducted in each study chapter. They have been released to foster mobile thermal imaging-empowered research.

In the following sections, we further discuss our contributions, their impacts and insights as well as future directions.

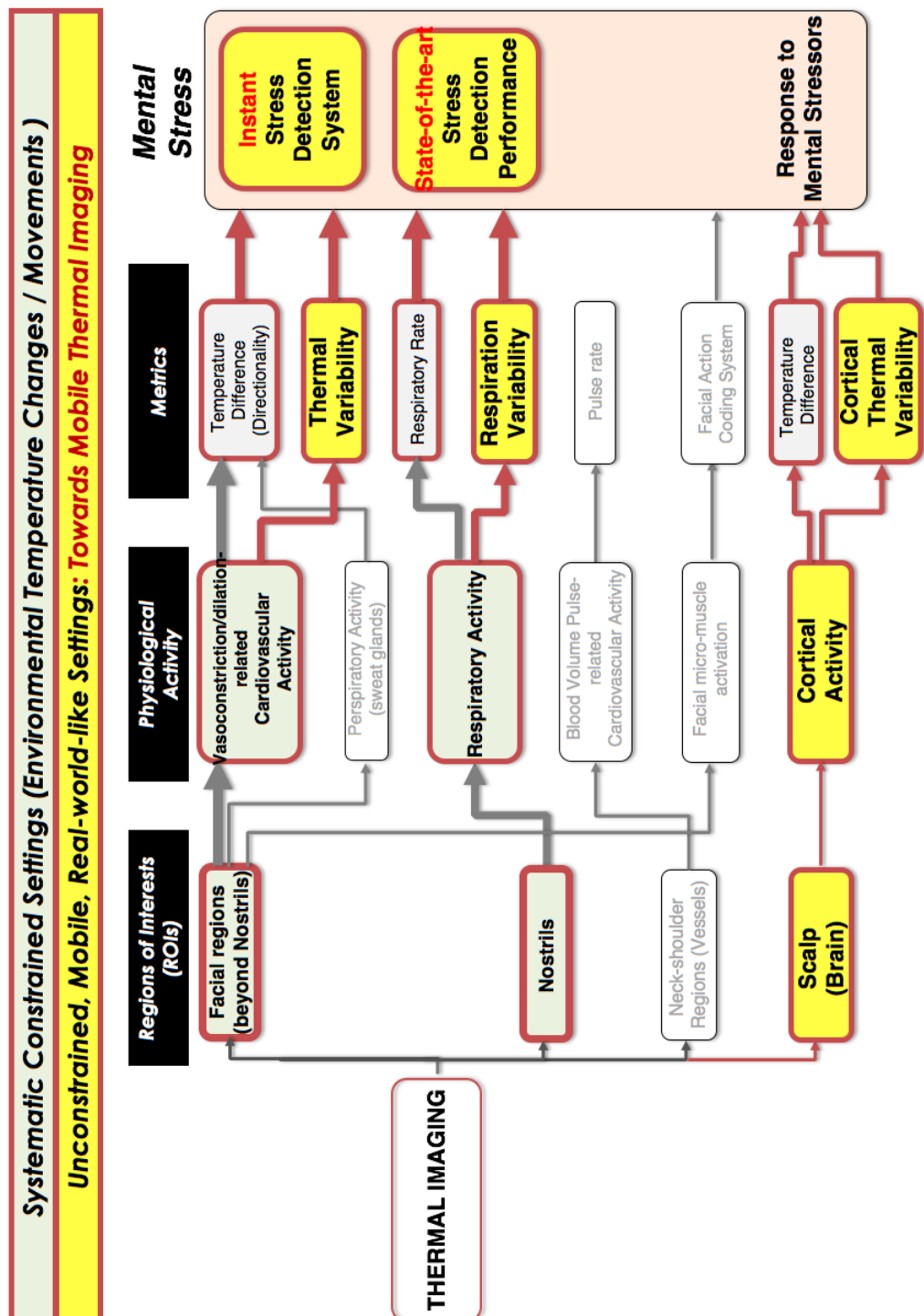


Figure 10.1. A summary diagram of key contributions of this thesis. Red bold lines (border, arrows) indicate what we have explored and improved with respect to state-of-the-arts. Yellow boxes with red borders highlight new perspectives of our novel contributions to this research area, filling the gaps on the diagram summarising the literature in Figure 2.2.

10.1. Mobile Thermal Imaging: Bringing Thermal Imaging into Unconstrained, Mobile, Real-world Settings

The ability to monitor physiological vital signs and psychological states is extremely important for medical treatment, healthcare and fitness sectors. In many situations, mobile methods, which allow users to undertake every day activities, are required. However, typically used technology for the monitoring can be obtrusive, requiring users to wear sensors such as chest belts, probes and electrodes (Healey & Picard, 2005; Hong *et al.*, 2012; Alberdi *et al.*, 2016). Alternatively, contactless digital image sensor based remote-photoplethysmography (PPG) can be used (Verkruysse *et al.*, 2008; Xu *et al.*, 2014; McDuff *et al.*, 2016; Gastel *et al.*, 2016). However, remote PPG requires an ambient source of light, and does not work properly in dark places or under varying lighting conditions. It also raises privacy issues. Thermal imaging, a non-contact method to measure heat patterns of materials and organisms, can be free from those constraints.

Initial evidence of possible use of thermal imaging in physiological and affective computing were discovered from earlier studies, for example, monitoring breathing rate (Murthy & Pavlidis, 2006; Lewis *et al.*, 2011; Pereira *et al.*, 2015) and affective states (Genno *et al.*, 1997; Puri *et al.*, 2005; Nhan & Chau, 2010; Engert *et al.*, 2014; Abdelrahman *et al.*, 2017). However, findings from their constrained, systematic laboratory settings are hard to generalise towards the real-world environment given constraints involved in the studies: controlled environmental temperature and one's mobility, as well as the use of high-cost and heavy thermographic systems. To tackle the issues and support one's real-life situations, we have proposed a new concept: ***Mobile Thermal Imaging***.

Recent advances in thermographic systems have shrunk their size, weight and cost, to the point where it is possible to create mobile monitoring devices that are not affected by lighting conditions. Although all our computational methods proposed in this thesis which achieved state-of-the-art performance are sensor-independent, we have mainly used low-cost, low-resolution thermal cameras (in Table 2.1, Figure 3.2) as main thermographic channels in this thesis. This provides two benefits. By doing so, we have been able to i) *measure temperature anywhere in mobile, temperature varying situations* and ii) *challenge our computational models given the unconstrained situations and low signal quality*. As we discuss details of the latter in the following sections, here we talk more about our design concepts for mobile thermal imaging explored in this thesis.

Proposed design concepts for mobile thermal imaging

In exploring mobile thermal imaging in unconstrained situations, we have presented two types of usage concepts: i) hardware-aided design (Figure 10.2), ii) non-hardware-aided design (Figure 10.3). As for the first design concept, the camera can be integrated into devices already used: rigid (e.g. PC, desk) or non-rigid (e.g. headset with microphone and Augmented Reality interface like HoloLens) structures to unobtrusively monitor a person's temperature. In Chapters 4, 5 and 7 we have built, and used, some of these physical structures, as shown in Figure 10.2, similar to set-ups used in other studies (Harrison *et al.*, 2011; Cho *et al.*, 2016b). To monitor a person's temperature while sitting on a chair (even wheelchairs), a hardware mount can be attached to place thermal cameras near a person's face, upper body (Figure 10.2a) and near the head (Figure 10.2b).

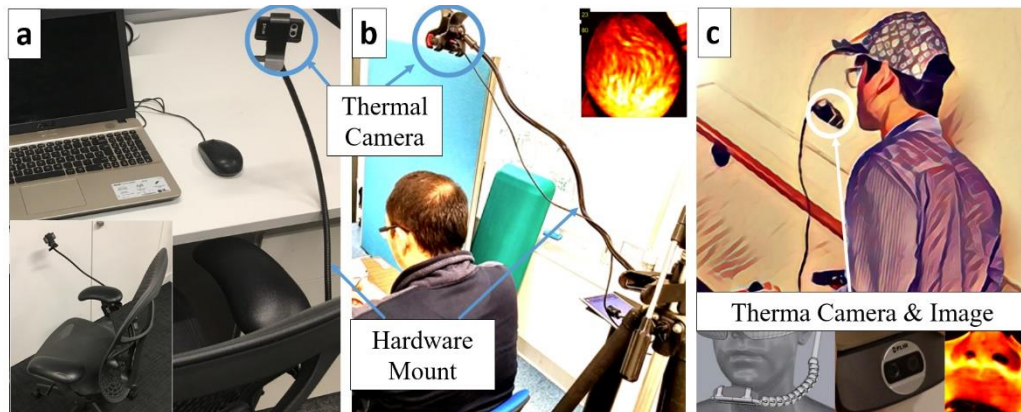


Figure 10.2. Design concept for mobile thermal imaging (hardware-aided design): a) a hardware frame attached to a chair for monitoring a person's upper body and face as in Chapter 4, b) the frame attached to a tripod for monitoring the head in Chapter 7, c) a head-set interface attached to a cap for the camera closely viewing a person's face in Chapter 5.

To closely monitor a person's facial thermal signatures while walking and moving, a headset-shaped interface shown in Figure 10.2c can be used. This can also be used together with a wearable device such as the Microsoft augmented reality device HoloLens. Even if not investigated in this thesis, with this setup as shown in Figure 10.3, our new approaches could possibly enable improved assessment of wellbeing of a worker in their work environment by monitoring multiple physiological cues together through the single thermal imaging channel. At

the same time, the worker can be informed of detected stress levels to help to tailor work schedules to his/her psychological needs. For this, multimodal feedback (e.g. tactile and auditory feedback as the author has explored beyond this thesis in Cho *et al.*, 2016a and Cho 2018) may also be used.

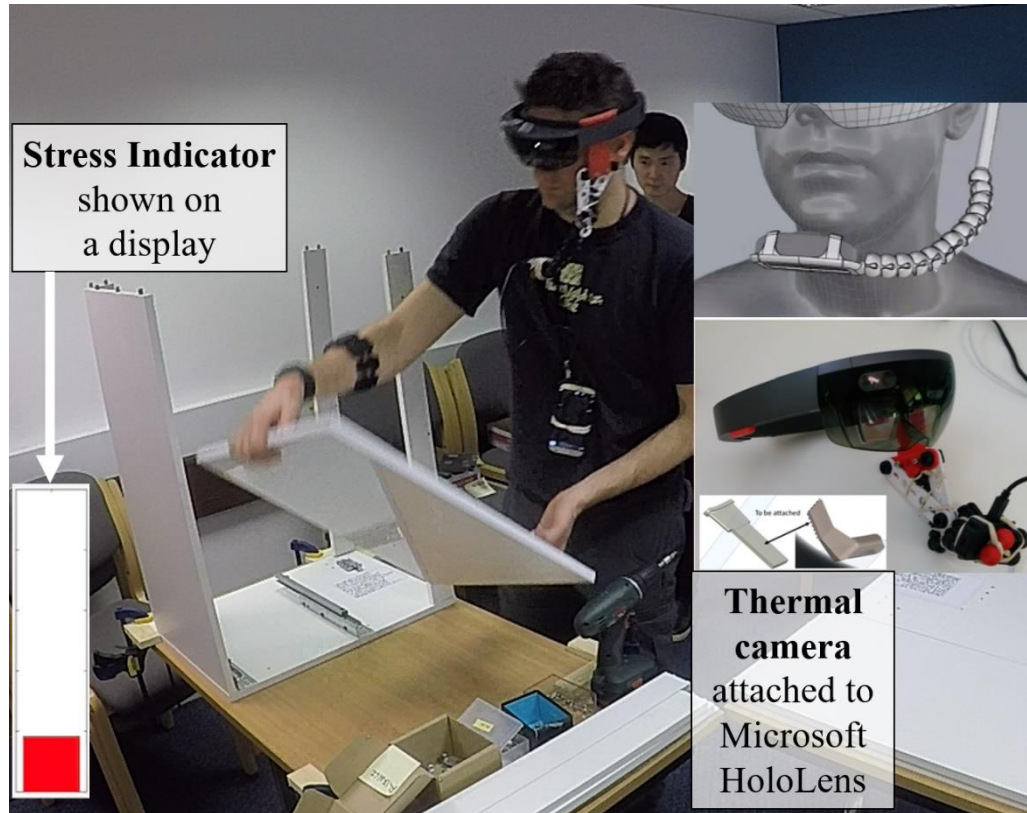


Figure 10.3. Mobile thermal imaging-based in-the-wild stress intervention and management for a human worker in a factory (the EU H2020 Human research project).

As alternative setups, Figure 10.4 describes non-hardware-aided design concepts. In particular, the handheld design used in Chapter 9 (and Figure 10.4a) provides more flexibility and scalability for recovering physiological signatures and recognising mental stress. Users can then use the system ad-hoc in situations when they need to monitor their condition. For example, a person can just hold a smartphone facing his/her face whilst walking to monitor breathing patterns, as shown in Figure 10.5. The visualised breathing pattern is produced by our proposed robust respiration tracking algorithm in Chapter 5. This flexible handheld setup enables many other uses, such as attaching the system on a desktop, desk, a lamp or handle of the bicycle, at a dashboard or even on a window in the car (Figure 10.4b, 10.4c, 10.4d). This flexible setup can

help thermographic cameras to effectively capture physiological thermal signatures of interest, whilst at the same time they do not interfere too much with a person's view.

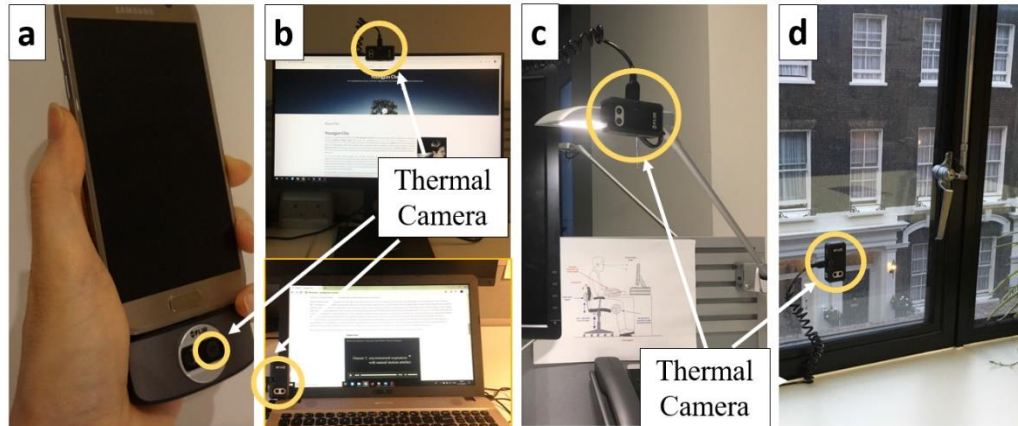


Figure 10.4. Design concept for mobile thermal imaging (without the aid of hardware): a) a handheld concept in Chapter 9, b) the camera can be also attached to, or near, a monitor as in Chapters 6, 8, alternatively, it can be simply attached to c) a lamp, even d) a window.

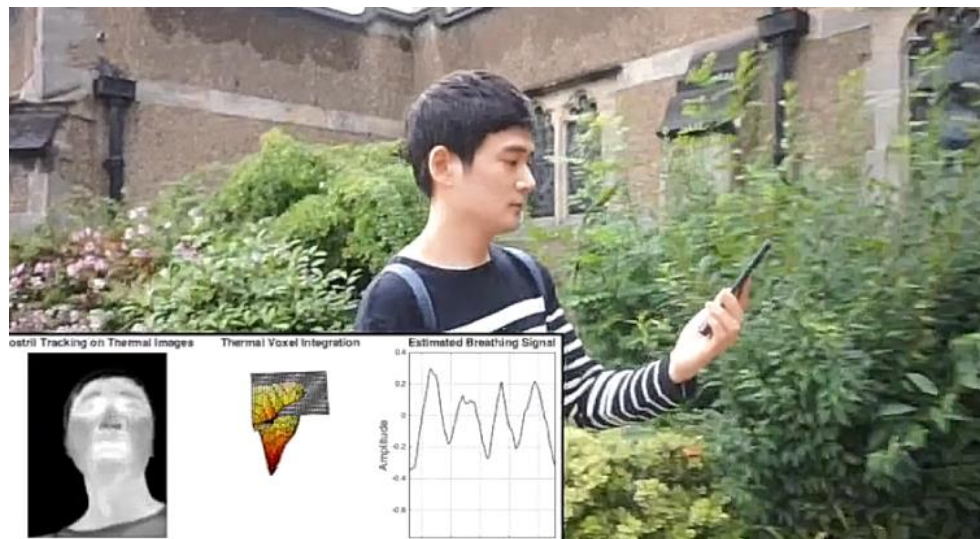


Figure 10.5. Mobile thermal imaging-powered anytime, anywhere with breathing monitoring (which is based upon our proposed robust respiration tracking algorithm in Chapter 5).

All in all, mobile thermography-based physiological computing and stress recognition, which this thesis has brought to the world for the first time, can perform in-the-wild settings, including outdoors and places under night illumination (i.e. key limitations of remote PPG, Verkrusse *et al.*, 2008), as well as removing the need to wear sensors (e.g. breathing belt) that

may not be suitable or unpleasant. For example, in the gym people may find sensors unpleasant due to sweat or constrained movements. Another case would be people with a certain type of pain such as CRPS (Birklein, 2005) who tend to reduce the number of clothes and objects that touch their skin. Given this, this new concept, mobile thermal imaging, could drive its scalability in real world settings.

10.2. ROI Tracking and Environmental Temperature Changes

To capacitate mobile thermal imaging to function, an important step is to automatically track Regions-Of-Interest (ROIs) on thermal images. In unconstrained contexts, this is inevitable to extract a certain type of physiological signature from each appropriate area (e.g. respiration: capturing heat exchanges in nostrils by inhalation and exhalation cycles of mammal breathing, Murthy & Pavlidis, 2006; Lewis *et al.*, 2011; Ebisch *et al.*, 2012). Some earlier studies conducted in highly constrained settings have also used ROI tracking algorithms originally built for RGB vision image sequences in order to improve performance in obtaining thermal signatures under motion artefacts (Sun *et al.*, 2005; Garbey *et al.*, 2007; Tsiamyrtzis *et al.*, 2007; Zhou *et al.*, 2009; Ebisch *et al.*, 2012; Shastri *et al.*, 2012; Manini *et al.*, 2013).

Beyond laboratory environments with stable temperature only, ambient temperature changes influence morphological and graphical properties of thermal images dynamically, negatively (e.g. Figure 4.1), making it harder to track ROIs. This has never been addressed in this area given the focus of the literature on static thermal imaging in highly controlled environments. To address this, in Chapter 4 we have proposed the Optimal Quantisation technique that adaptively quantises temperature values to complement environmental thermal dynamics on thermal images. We have identified how use of the method alone can significantly enhance performance of existing algorithms in tracking ROIs (e.g. periorbital area in Table 4.1) on thermal imaging sequences even in indoor laboratory settings with varying environmental temperatures in Chapter 4. We have used this technique as the backbone to build our computational methods proposed in other chapters for robust tracking of physiological signals and automatically detecting mental stress in more unconstrained settings. For example, with our nostril-ROI tracking algorithm proposed in Chapter 5, the Optimal Quantisation has achieved extreme robustness in tracking the area even during the outdoor physical exercise session.

Furthermore, we have also shown how positively this method can generalise to other material types than the human skin. In other mobile, ubiquitous applications using thermal imaging beyond physiological and affective computing, we have found the significant role of the quantisation process. With a deep learning framework (transformer network-based convolutional neural network; Jaderberg *et al.*, 2015) shown together in Chapter 4, dynamically quantising a thermal matrix makes it possible to learn spatial thermal patterns of material surfaces outdoor so computing devices can recognise their types. We believe this fascinating capability can boost the general use of thermal imaging in enhancing our experiences within many HCI settings.

10.3. Representation of Physiological Thermal Signature

How to represent signals is an important question in physiological computing to obtain reliable physiological cues. Averaging has been the most common, simplest way to represent a physiological signal. For example, averaging temperatures within a nostril ROI (usually, rectangular shape) has been most frequently used to represent a breathing cue on a thermal image (Murthy & Pavlidis, 2006; Fei & Pavlidis, 2010; Abbas *et al.*, 2011; Lewis *et al.*, 2011; Pereira *et al.*, 2015). The same applies to capturing thermal directionality which has been a main focus in the literature (Veltman & Vos, 2005; Or & Duffy, 2007; Di Giacinto *et al.*, 2014; Engert *et al.*, 2014; Abdelrahman *et al.*, 2017). Beyond the main scope of this thesis, methods that enable a camera-based PPG are all based on this averaging technique (averaging intensity values on images as explained in Appendix F.1).

However, in mobile, unconstrained, outdoor situations, it is hard to ensure the acquisition of a high quality of this certain type of physiological thermal signature, in particular, through mobile thermal imaging (even if an enhanced quantisation process adjusts environmental thermal effects). For example, a person's shallow breathing would accompany very subtle heat exchanges, resulting in blurred patterns on thermal images (e.g. Figure 5.1b). In this case, averaging is likely to lose the fairly small but important information, as discussed in Chapter 5. Towards this end, we have introduced a novel signal representation based on the new concept of Thermal Voxel (Section 5.2.2), which helps to obtain a higher quality of respiratory signals in comparison with the traditional method (i.e. average temperature), as shown in Figures 5.10-5.14.

Likewise, in Chapter 7, we have also identified the same need for improved capturing of a cortical physiological cue from a spatial thermal pattern, rather than averaging. Exploring novel stress related cortical thermal signatures from the unexplored area, the scalp, we have shown that a timeseries of spatial skew helps to capture cortical activation/deactivation-induced thermal variations over the scalp whilst averaging does not. In turn, our novel signal representation techniques have enhanced signal quality of physiological patterns extracted through thermal imaging, providing temporally richer information where we can capture physiological variability.

10.4. Important Role of Tracking Physiological Variability in Stress Assessment

Extreme robustness of physiological measurements has to be made in order to reliably, automatically assess a person's physiological and psychological states, in particular, mental stress. Robust methods which this thesis has contributed for extracting respiratory patterns (Chapter 5) and vasoconstriction/dilation-induced nasal temperature variation patterns (Chapter 6), can produce richer, less noisy physiological signals than existing methods do. This can play a significant role in automated stress inference. By richness, we are referring to the complexity of physiological signatures.

As for the respiratory signature, our method (Chapter 5) can capture variations along with respiratory cycles. Likewise, as for the cardiovascular signature from the nose tip area, our method (Chapter 6) can capture complex variations along with vasoconstriction and vasodilation patterns. This can help to explore variability beyond single metric values, such as a simplistic respiratory rate or a binary thermal direction which has been primarily explored in the literature. In particular, studies where thermal imaging is used for affective computing have mostly focused on exploring the relationship between the binary direction of thermal changes on facial or palm areas and affective states (Genno *et al.*, 1997; Pavlidis *et al.*, 2001; Puri *et al.*, 2005; Gane *et al.*, 2011; Engert *et al.*, 2014). In addition, the use of the single metric has resulted in inconsistent findings (e.g. Pavlidis *et al.*, 2001 and Gane *et al.*, 2011).

As discussed in Chapter 6, the multiple metrics which are designed to capture variations helped produce significance while the existing single metric did not in quantifying mental stress. Given this, the richer physiological information gathered by our methods can now empower the

exploration of the role of each physiological thermal signature in mental stress recognition in Chapters 8 and 9 where the contribution of each signature to automatic stress inference has been separately investigated (Chapter 8: revisiting the role of breathing in existing stress inference protocols, Chapter 9: exploring the role of nasal thermal variability in instant stress recognition).

The literature of mental stress recognition or quantification (including thermal imaging-based studies) has mostly focused on gross statistical features. For example, in the case of breathing, average breathing rate over a specified time window (e.g. 2 minutes) has been used as the main feature (Hong *et al.*, 2012; McDuff *et al.*, 2016). However, given that mental stress induces complex physiological phenomenon (Everly Jr & Lating, 2012), the contribution of such gross statistical feature has been very weak or unclear towards mental stress recognition (McDuff *et al.*, 2016; Masaoka & Homma, 1997; Hong *et al.*, 2012). Inspired by heart rate variability metrics (e.g. LF/HF ratio, SDNN) that help assess sympathovagal balance between the SNS and PSNS (Billman, 2013; Shaffer & Ginsberg, 2017), we have focused on how to represent variability of the rich respiratory (Chapter 8) and vasoconstriction/dilation-induced nasal thermal signatures (Chapter 9).

In particular, with the case of respiration, stressful situations induce irregular breathing patterns, according to Grossman (1983). This may be due to our fight-or-flight responses of the ANS to stressors (Everly Jr & Lating, 2012). Given this, we have proposed a new method with the use of bi-dimensional spectrogram to capture variability of breathing patterns in a compact way, considering both the time and frequency domains, discussed in Chapter 8. Following this, we have also proposed the use of automatic feature learning (representation learning) using artificial neural networks (e.g. deep convolutional neural network, LeCun *et al.*, 2015) to automatically find good features in relation to mental stress during the learning process from the newly represented low-level feature (i.e. Respiration Variability Spectrogram, RVS).

Likewise, we have also proposed to use representation learning techniques by taking the uni-dimensional nasal thermal variability sequence (Figure 9.3) from a short-term measurement as low-level features in Chapter 9. This strategy can eliminate the need for feature handcrafting (i.e. carefully designing metrics and use them as machine learning features). As discussed in LeCun *et al.* (2015), even carefully hand engineered-feature extractors could fail to generalise towards unseen data sets. Throughout the conducted evaluation study for each approach, we have demonstrated that i) a new representation of breathing variability has enabled to achieve the state-of-the-art stress recognition performance from the use of 2 minutes-breathing signals only (accuracy of 84.59%, LOSO validation) and ii) the use of the new nasal cardiovascular thermal

variability together with blood volume pulse from another channel has helped achieve an accuracy of 78.33% from a very short term measurement (20 seconds, we call this *instant stress detection*), which is also the latest performance.

10.5. Limitation and Future Directions

There is still room for improvement towards building reliable physiological monitoring and mental stress recognition systems. Despite the improved quantisation process which helps to address effects of environmental temperature on thermal imaging of a person's facial or scalp morphological shape, the ambient temperature still influences one's physiological pattern. For example, from the study for instant stress inference in Chapter 9, data in Rest 1 was affected by the difference in temperature between the waiting area and experiment area. This effect was further enhanced when the participants had just arrived from outdoors with the body temperature being strongly influenced by colder outdoor weather (winter season). This is important, as if the system has to be used, it is crucial that the person uses it in the same environment where the stressor occurs. It should also be tested in future studies if the decrease in nose tip temperature may be saturated by very cold environment temperatures and therefore would be less informative in such situations for stress detection. For real world applications, it is critical to take into account such factors in signal processing or to inform users that the system is not reliable in the case of saturation.

Furthermore, there is a need to explore how this mobile thermal imaging-enabled physiological computing and stress recognition capability can help manage stress. Indeed, smartphone apps with such capabilities are increasingly desired as possible tools for facilitating stress self-management (Konrad et al., 2015; Coulon et al., 2016; Ptakauskaite et al., 2018), as people are often unaware of their level of stress as well as being stress-sensitive to particular situations; for example, chronic pain can cause a fear of movement (Felipe et al., 2015). There is also strong interest within the industry to complement typically used questionnaires which enable improved assessment of wellbeing with personnel, as well as revisiting work plans and work environments (Fleck et al., 2015). As discussed in Section 10.1, mobile thermal imaging sensing channels could be embedded into employees' aids for ease of use. While these low-cost sensors are still not perfect, we have contributed towards enhancing their capability. At the same time,

we hope our work contributes to the literature in general using these signals as a measure of mental stress (Mohan et al., 2016; Jobbágy et al., 2017; Charlton et al., 2018).

Although our main aim for Stage III of this thesis was to build representation techniques that can enhance the stress inference performance of existing machine learning classifiers, we need to note the limitations of machine learning algorithms we used, in particular, convolutional neural networks. Given that the model has a large number of parameters, it is generally susceptible to overfitting, leading to poor performance on unseen data. The dropout technique can help address this issue as discussed in Section 4.6 (Srivastava *et al.*, 2014). In addition, CNNs are not ideal to handle physiological signals that are time-varying rather than just spatially variant. Hence, in future work, we plan to implement further machine learning algorithms which are to process sequential data, such as, a Recurrent Neural Network (RNN). In addition, attention mechanisms could be used to overcome the limitation of CNN. Recent literature has shown that CNN may become as powerful as RNN if such mechanisms are used. In parallel, we also plan to merge different classifiers into one decision model so as to fuse stress-related information from multiple thermal signatures.

Lastly, we have found that constraints arise from the resolution and imaging quality of low-cost thermal imaging systems themselves. In particular, low spatial resolution limits the distance at which physiological patterns can be measured (in our conducted data collection studies, less than 0.7 meters). In addition, the thermal sensitivity ($<0.1^{\circ}\text{C}$) was also limited, which may interfere with capturing subtle physiological patterns (e.g. cortical thermal signatures from the scalp – there may be more subtle, delicate cortical signals as signals EEG can detect). When we first undertook this work, only the FLIR One (2G) was available as a mobile thermographic channel. However, now we can see more options coming to the market such as Therm-App Hz (see Table 2.1). Given high performance produced by our proposed methods, we can expect that by using new low-cost, but more thermal sensitive cameras which are emerging, their performance will get even better. Indeed, most of our proposed solutions are camera-parameter independent and therefore may lead to further improvements with enhanced camera properties. Higher quality mobile thermal imaging will to a greater extent open up exciting prospects for physiological and affective computing, as well as HCI.

References

- Abbas AK, Heimann K, Jergus K, Orlikowsky T & Leonhardt S (2011). Neonatal non-contact respiratory monitoring based on real-time infrared thermography. *Biomed Eng OnLine* **10**, 93–109.
- Abdelrahman Y, Velloso E, Dingler T, Schmidt A & Vetere F (2017). Cognitive heat: exploring the usage of thermal imaging to unobtrusively estimate cognitive load. *Proc ACM Interact Mob Wearable Ubiquitous Technol* **1**, 33.
- Abouelenien M, Pérez-Rosas V, Mihalcea R & Burzo M (2017). Detecting Deceptive Behavior via Integration of Discriminative Features From Multiple Modalities. *IEEE Trans Inf Forensics Secur* **12**, 1042–1055.
- Aittala M, Weyrich T & Lehtinen J (2015). Two-shot SVBRDF capture for stationary materials. *ACM Trans Graph TOG - Proc ACM SIGGRAPH 2015* **34**, 110–122.
- Åkerstedt T, Gillberg M, Hjemdahl P, Sigurdson K, Gustavsson I, Daleskog M & Pollare T (1983). Comparison of urinary and plasma catecholamine responses to mental stress. *Acta Physiol Scand* **117**, 19–26.
- Alberdi A, Aztiria A & Basarab A (2016). Towards an automatic early stress recognition system for office environments based on multimodal measurements: A review. *J Biomed Inform* **59**, 49–75.
- Allen J (2007). Photoplethysmography and its application in clinical physiological measurement. *Physiol Meas* **28**, R1.
- Al-Shargie F, Kiguchi M, Badruddin N, Dass SC, Hani AFM & Tang TB (2016). Mental stress assessment using simultaneous measurement of EEG and fNIRS. *Biomed Opt Express* **7**, 3882–3898.
- Altman DG & Bland JM (1983). Measurement in Medicine: The Analysis of Method Comparison Studies. *J R Stat Soc Ser Stat* **32**, 307–317.
- Anonymous (1877). Cerebral Thermometry. *The Lancet* **110**, 586.
- Arnsten AFT (2009). Stress signalling pathways that impair prefrontal cortex structure and function. *Nat Rev Neurosci* **10**, 410–422.
- Ashburner J & Friston KJ (2000). Voxel-Based Morphometry—The Methods. *NeuroImage* **11**, 805–821.

- Atkinson D, Baurley S, Petreca BB, Bianchi-Berthouze N & Watkins P (2016). The Tactile Triangle: a design research framework demonstrated through tactile comparisons of textile materials. *J Des Res* **14**, 142–170.
- Avants BB, Epstein CL, Grossman M & Gee JC (2008). Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Med Image Anal* **12**, 26–41.
- Avants BB, Tustison NJ, Song G, Cook PA, Klein A & Gee JC (2011). A reproducible evaluation of ANTs similarity metric performance in brain image registration. *NeuroImage* **54**, 2033–2044.
- Bang AL (2009). Triads as a means for dialogue about emotional values in textile design. In *Proceedings of: 8th European Academy of Design Conference: Design Connexity*. Citeseer.
- Barclay CJ, Arnold PD & Gibbs CL (1995). Fatigue and heat production in repeated contractions of mouse skeletal muscle. *J Physiol* **488**, 741–752.
- Bartuzi P, Roman-Liu D & Wiśniewski T (2012). The Influence of Fatigue on Muscle Temperature. *Int J Occup Saf Ergon* **18**, 233–243.
- Bebis G, Gyaourova A, Singh S & Pavlidis I (2006). Face recognition by fusing thermal infrared and visible imagery. *Image Vis Comput* **24**, 727–742.
- Bell S, Upchurch P, Snavely N & Bala K (2015). Material Recognition in the Wild With the Materials in Context Database. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3479–3487.
- Belujon P & Grace AA (2015). Regulation of dopamine system responsivity and its adaptive and pathological response to stress. *Proc R Soc B* **282**, 20142516.
- Bernardi L, Wdowczyk-Szulc J, Valenti C, Castoldi S, Passino C, Spadacini G & Sleight P (2000). Effects of controlled breathing, mental activity and mental stress with or without verbalization on heart rate variability. *J Am Coll Cardiol* **35**, 1462–1469.
- Bijur PE, Silver W & Gallagher EJ (2001). Reliability of the Visual Analog Scale for Measurement of Acute Pain. *Acad Emerg Med* **8**, 1153–1157.
- Billman GE (2011). Heart Rate Variability – A Historical Perspective. *Front Physiol*; DOI: 10.3389/fphys.2011.00086.
- Billman GE (2013). The LF/HF ratio does not accurately measure cardiac sympatho-vagal balance. *Front Physiol*; DOI: 10.3389/fphys.2013.00026.
- Birklein F (2005). Complex regional pain syndrome. *J Neurol* **252**, 131–138.
- Boto E, Holmes N, Leggett J, Roberts G, Shah V, Meyer SS, Muñoz LD, Mullinger KJ, Tierney TM, Bestmann S, Barnes GR, Bowtell R & Brookes MJ (2018). Moving magnetoencephalography towards real-world applications with a wearable system. *Nature* **555**, 657–661.

- Brosschot JF, Gerin W & Thayer JF (2006). The perseverative cognition hypothesis: A review of worry, prolonged stress-related physiological activation, and health. *J Psychosom Res* **60**, 113–124.
- Brüser C, Winter S & Leonhardt S (2013). Robust inter-beat interval estimation in cardiac vibration signals. *Physiol Meas* **34**, 123–138.
- Busto R, Dietrich WD, Globus MY-T, Valdés I, Scheinberg P & Ginsberg MD (1987). Small Differences in Intrascemic Brain Temperature Critically Determine the Extent of Ischemic Neuronal Injury. *J Cereb Blood Flow Metab* **7**, 729–738.
- Byng JW, Boyd NF, Fishell E, Jong RA & Yaffe MJ (1996). Automated analysis of mammographic densities. *Phys Med Biol* **41**, 909.
- Caffier PP, Erdmann U & Ullsperger P (2003). Experimental evaluation of eye-blink parameters as a drowsiness measure. *Eur J Appl Physiol* **89**, 319–325.
- Calatayud J, Borreani S, Colado JC, Martín FF, Rogers ME, Behm DG & Andersen LL (2014). Muscle Activation during Push-Ups with Different Suspension Training Systems. *J Sports Sci Med* **13**, 502–510.
- Callister WD & Rethwisch DG (2011). *Materials science and engineering: an introduction*. John Wiley & Sons NY.
- Camm AJ, Malik M, Bigger JT, Breithardt G, Cerutti S, Cohen RJ, Coumel P, Fallen EL, Kennedy HL, Kleiger R, Lombardi F, Malliani A, Moss AJ, Rottman JN, Schmidt G, Schwartz PJ & Singer D (1996). Heart rate variability. Standards of measurement, physiological interpretation, and clinical use. *Eur HEART J* **17**, 354–381.
- Causse M, Chua Z, Peysakhovich V, Campo ND & Matton N (2017). Mental workload and neural efficiency quantified in the prefrontal cortex using fNIRS. *Sci Rep* **7**, 5222.
- Chan P-H, Wong C-K, Poh YC, Pun L, Leung WW-C, Wong Y-F, Wong MM-Y, Poh M-Z, Chu DW-S & Siu C-W (2016). Diagnostic performance of a smartphone-based photoplethysmographic application for atrial fibrillation screening in a primary care setting. *J Am Heart Assoc* **5**, e003428.
- Charlton PH, Celka P, Farukh B, Chowienczyk P & Alastruey J (2018). Assessing mental stress from the photoplethysmogram: a numerical study. *Physiol Meas* **39**, 054001.
- Cheng Y (1995). Mean shift, mode seeking, and clustering. *IEEE Trans Pattern Anal Mach Intell* **17**, 790–799.
- Cho Y (2018). Sensorless Resonance Tracking of Resonant Electromagnetic Actuator through Back-EMF Estimation for Mobile Devices. *ArXiv Prepr ArXiv180307065*.
- Cho Y, Bianchi A, Marquardt N & Bianchi-Berthouze N (2016a). RealPen: Providing Realism in Handwriting Tasks on Touch Surfaces Using Auditory-Tactile

- Feedback. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, pp. 195–205. ACM, New York, NY, USA. Available at: <http://doi.acm.org/10.1145/2984511.2984550>.
- Cho Y, Bianchi-Berthouze N & Julier SJ (2017a). DeepBreath: Deep Learning of Breathing Patterns for Automatic Stress Recognition using Low-Cost Thermal Imaging in Unconstrained Settings. In *the 7th International Conference on Affective Computing and Intelligent Interaction, ACII 2017*, pp. 456–463.
- Cho Y, Bianchi-Berthouze N, Julier SJ & Marquardt N (2017b). ThermSense: Smartphone-based breathing sensing platform using noncontact low-cost thermal camera. In *2017 Seventh International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)*, pp. 83–84.
- Cho Y, Bianchi-Berthouze N, Marquardt N & Julier SJ (2018). Deep Thermal Imaging: Proximate Material Type Recognition in the Wild Through Deep Learning of Spatial Surface Temperature Patterns. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 2:1–2:13. ACM, New York, NY, USA. Available at: <http://doi.acm.org/10.1145/3173574.3173576>.
- Cho Y, Julier SJ & Bianchi-Berthouze N (2019). Instant Stress: Detection of Perceived Mental Stress through Smartphone Photoplethysmography and Thermal Imaging. *JMIR Ment Health* (in press). DOI: 10.2196/10140.
- Cho Y, Julier SJ, Marquardt N & Bianchi-Berthouze N (2017c). Robust tracking of respiratory rate in high-dynamic range scenes using mobile thermal imaging. *Biomed Opt Express* **8**, 4480–4503.
- Cho Y, Kim S & Joung M (2016b). Proximity Sensor and Control Method Thereof. US Patent, patent number: 20160061588.
- Cimpoi M, Maji S & Vedaldi A (2015). Deep Filter Banks for Texture Recognition and Segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3828–3836.
- Clifford GD, Behar J, Li Q & Rezek I (2012). Signal quality indices and data fusion for determining clinical acceptability of electrocardiograms. *Physiol Meas* **33**, 1419.
- Costafreda SG, Fu CHY, Lee L, Everitt B, Brammer MJ & David AS (2006). A systematic review and quantitative appraisal of fMRI studies of verbal fluency: Role of the left inferior frontal gyrus. *Hum Brain Mapp* **27**, 799–810.
- Coulon SM, Monroe CM & West DS (2016). A systematic, multi-domain review of mobile smartphone apps for evidence-based stress management. *Am J Prev Med* **51**, 95–105.
- Crawford A (1788). *Experiments and observations on animal heat and the inflammation of combustible bodies*. Johnson.

- Cretikos MA, Bellomo R, Hillman K, Chen J, Finfer S & Flabouris A (2008). Respiratory rate: the neglected vital sign. *Med J Aust* **188**, 657–659.
- Dedovic K, D’Aguiar C & Pruessner JC (2009). What Stress Does to Your Brain: A Review of Neuroimaging Studies. *Can J Psychiatry* **54**, 6–15.
- Dedovic K, Renwick R, Mahani NK, Engert V, Lupien SJ & Pruessner JC (2005). The Montreal Imaging Stress Task: using functional imaging to investigate the effects of perceiving and processing psychosocial stress in the human brain. *J Psychiatry Neurosci* **30**, 319–326.
- Dey AK & Häkkinen J (2008). Context-Awareness and Mobile Devices. *Handb Res User Interface Des Eval Mob Technol* 205–217.
- Di Giacinto A, Brunetti M, Sepede G, Ferretti A & Merla A (2014). Thermal signature of fear conditioning in mild post traumatic stress disorder. *Neuroscience* **266**, 216–223.
- Donkin H (1879). On Some Cases of Abnormally High Temperature. *Br Med J* **2**, 983–984.
- Dourish P (2004). What we talk about when we talk about context. *Pers Ubiquitous Comput* **8**, 19–30.
- Droitcour AD, Boric-Lubecke O & Kovacs GTA (2009). Signal-to-Noise Ratio in Doppler Radar System for Heart and Respiratory Rate Measurements. *IEEE Trans Microw Theory Tech* **57**, 2498–2507.
- Duda RO & Hart PE (1972). Use of the Hough transformation to detect lines and curves in pictures. *Commun ACM* **15**, 11–15.
- Dunn AJ & Berridge CW (1990). Physiological and behavioral responses to corticotropin-releasing factor administration: is CRF a mediator of anxiety or stress responses? *Brain Res Brain Res Rev* **15**, 71–100.
- Ebisch SJ, Aureli T, Bafunno D, Cardone D, Romani GL & Merla A (2012). Mother and child in synchrony: Thermal facial imprints of autonomic contagion. *Biol Psychol* **89**, 123–129.
- Ekman P (1993). Facial expression and emotion. *Am Psychol* **48**, 384–392.
- Ekman P & Friesen W (1977). *Facial Action Coding System*. Consulting Psychologists Press, Stanford University, Palo Alto.
- Elam M & Wallin BG (1987). Skin blood flow responses to mental stress in man depend on body temperature. *Acta Physiol Scand* **129**, 429–431.
- Elgendi M (2016). Optimal Signal Quality Index for Photoplethysmogram Signals. *Bioengineering*; DOI: 10.3390/bioengineering3040021.

- Engert V, Merla A, Grant JA, Cardone D, Tusche A & Singer T (2014). Exploring the Use of Thermal Infrared Imaging in Human Stress Research. *PLOS ONE* **9**, e90782.
- Everly Jr GS & Lating JM (2012). *A clinical guide to the treatment of the human stress response*. Springer Science & Business Media.
- Fahrenheit DG (1724). Experimenta & Observationes De Congelatione Aquae in Vacuo Factae a D. G. Fahrenheit, R. S. S. *Philos Trans 1683-1775* **33**, 78–84.
- Fei J & Pavlidis I (2010). Thermistor at a Distance: Unobtrusive Measurement of Breathing. *IEEE Trans Biomed Eng* **57**, 988–998.
- Felipe S, Singh A, Bradley C, Williams AC & Bianchi-Berthouze N (2015). Roles for Personal Informatics in chronic pain. In *2015 9th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth)*, pp. 161–168.
- Fleck R, Cox AL & Robison RAV (2015). Balancing Boundaries: Using Multiple Devices to Manage Work-Life Balance. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pp. 3985–3988. <http://doi.acm.org/10.1145/2702123.2702386>.
- Gane L, Power S, Kushki A & Chau T (2011). Thermal Imaging of the Periorbital Regions during the Presentation of an Auditory Startle Stimulus. *PLOS ONE* **6**, e27268.
- Garbey M, Sun N, Merla A & Pavlidis I (2007). Contact-Free Measurement of Cardiac Pulse Based on the Analysis of Thermal Imagery. *IEEE Trans Biomed Eng* **54**, 1418–1426.
- Gastel M van, Stuijk S & Haan G de (2016). Robust respiration detection from remote photoplethysmography. *Biomed Opt Express* **7**, 4941–4957.
- Genno H, Ishikawa K, Kanbara O, Kikumoto M, Fujiwara Y, Suzuki R & Osumi M (1997). Using facial skin temperature to objectively evaluate sensations. *Int J Ind Ergon* **19**, 161–171.
- Ghosh S, Laksana E, Scherer S & Morency LP (2015). A multi-label convolutional neural network approach to cross-domain action unit detection. In *2015 International Conference on Affective Computing and Intelligent Interaction (ACII)*, pp. 609–615.
- Giardino ND, Lehrer PM & Edelberg R (2002). Comparison of finger plethysmograph to ECG in the measurement of heart rate variability. *Psychophysiology* **39**, 246–253.
- Girden ER (1992). *ANOVA: Repeated Measures*. SAGE.
- Govier AV, Reves JG, McKay RD, Karp RB, Zorn GL, Morawetz RB, Smith LR, Adams M & Freeman AM (1984). Factors and their influence on regional

- cerebral blood flow during nonpulsatile cardiopulmonary bypass. *Ann Thorac Surg* **38**, 592–600.
- Grant R (1950). Emotional hypothermia in rabbits. *Am J Physiol-Leg Content* **160**, 285–290.
- Gray RM & Davisson LD (2004). *An Introduction to Statistical Signal Processing*. Cambridge University Press.
- Grossman P (1983). Respiration, Stress, and Cardiovascular Function. *Psychophysiology* **20**, 284–300.
- Gur RC, Gur RE, Skolnick BE, Resnick SM, Silver FL, Chawluk J, Muenz L, Obrist WD & Reivich M (1988). Effects of Task Difficulty on Regional Cerebral Blood Flow: Relationships with Anxiety and Performance. *Psychophysiology* **25**, 392–399.
- Hahn AC, Whitehead RD, Albrecht M, Lefevre CE & Perrett DI (2012). Hot or not? Thermal reactions to social contact. *Biol Lett* **20120338**.
- Hall JE & Guyton AC (2015). *Guyton and Hall Textbook of Medical Physiology*. Elsevier Health Sciences.
- Haller Jr JS (1985). Medical thermometry—a short history. *West J Med* **142**, 108.
- Hamilton PS & Tompkins WJ (1986). Quantitative Investigation of QRS Detection Rules Using the MIT/BIH Arrhythmia Database. *IEEE Trans Biomed Eng* **BME-33**, 1157–1165.
- Harrison C, Benko H & Wilson AD (2011). OmniTouch: Wearable Multitouch Interaction Everywhere. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology*, pp. 441–450. ACM, New York, NY, USA. Available at: <http://doi.acm.org/10.1145/2047196.2047255>.
- Healey JA & Picard RW (2005). Detecting stress during real-world driving tasks using physiological sensors. *IEEE Trans Intell Transp Syst* **6**, 156–166.
- Heathers JAJ (2013). Smartphone-enabled pulse rate variability: An alternative methodology for the collection of heart rate variability in psychophysiological research. *Int J Psychophysiol* **89**, 297–304.
- Heinrichs SC, Li DL & Iyengar S (2001). Corticotropin-releasing factor (CRF) or CRF binding-protein ligand inhibitor administration suppresses food intake in mice and elevates body temperature in rats. *Brain Res* **900**, 177–185.
- Heo J, Kong SG, Abidi BR & Abidi MA (2004). Fusion of Visual and Thermal Signatures with Eyeglass Removal for Robust Face Recognition. In *Conference on Computer Vision and Pattern Recognition Workshop, 2004. CVPRW '04*, pp. 122–122.

- Hernández B, Olague G, Hammoud R, Trujillo L & Romero E (2007). Visual learning of texture descriptors for facial expression recognition in thermal imagery. *Comput Vis Image Underst* **106**, 258–269.
- Hernandez J, Morris RR & Picard RW (2011). Call Center Stress Recognition with Person-Specific Models. In *Affective Computing and Intelligent Interaction*, pp. 125–134. Springer, Berlin, Heidelberg. Available at: https://link.springer.com/chapter/10.1007/978-3-642-24600-5_16.
- Hjortskov N, Rissén D, Blangsted AK, Fallentin N, Lundberg U & Sjøgaard K (2004). The effect of mental stress on heart rate variability and blood pressure during computer work. *Eur J Appl Physiol* **92**, 84–89.
- Hong J-H, Ramos J & Dey AK (2012). Understanding Physiological Responses to Stressors During Physical Activity. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, pp. 270–279. ACM, New York, NY, USA. Available at: <http://doi.acm.org/10.1145/2370216.2370260>.
- Hosseini SA & Khalilzadeh MA (2010). Emotional Stress Recognition System Using EEG and Psychophysiological Signals: Using New Labelling Process of EEG Signals in Emotional Stress State. In *2010 International Conference on Biomedical Engineering and Computer Science*, pp. 1–6.
- Hovsepian K, al’Absi M, Ertin E, Kamarck T, Nakajima M & Kumar S (2015). cStress: Towards a Gold Standard for Continuous Stress Assessment in the Mobile Environment. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pp. 493–504. ACM, New York, NY, USA. Available at: <http://doi.acm.org/10.1145/2750858.2807526>.
- Ioannou S, Gallese V & Merla A (2014). Thermal infrared imaging in psychophysiology: Potentialities and limits. *Psychophysiology* **51**, 951–963.
- Isard M & Blake A (1998). CONDENSATION—Conditional Density Propagation for Visual Tracking. *Int J Comput Vis* **29**, 5–28.
- Jackson J (1828). Case of a Cartilaginous Tumor in the Trachea, Producing Dyspnoea, &c. *Boston Med Surg J* **1**, 33–36.
- Jaderberg M, Simonyan K, Zisserman A & kavukcuoglu koray (2015). Spatial Transformer Networks. In *Advances in Neural Information Processing Systems* 28, ed. Cortes C, Lawrence ND, Lee DD, Sugiyama M & Garnett R, pp. 2017–2025. Curran Associates, Inc.
- James W (1884). What is an Emotion? *Mind* **9**, 188–205.
- Jarlier S, Grandjean D, Delplanque S, N’Diaye K, Cayeux I, Velazco MI, Sander D, Vuilleumier P & Scherer KR (2011). Thermal Analysis of Facial Muscles Contractions. *IEEE Trans Affect Comput* **2**, 2–9.

- Jobbágy Á, Majnár M, Tóth LK & Nagy P (2017). HRV-based Stress Level Assessment Using Very Short Recordings. *Period Polytech Electr Eng Comput Sci* **61**, 238–245.
- Jonathan E & Leahy M (2010). Investigating a smartphone imaging unit for photoplethysmography. *Physiol Meas* **31**, N79.
- Kalal Z, Mikolajczyk K & Matas J (2010). Forward-Backward Error: Automatic Detection of Tracking Failures. In *2010 20th International Conference on Pattern Recognition (ICPR)*, pp. 2756–2759.
- Kalal Z, Mikolajczyk K & Matas J (2012). Tracking-Learning-Detection. *IEEE Trans Pattern Anal Mach Intell* **34**, 1409–1422.
- Karjalainen J, Viitasalo M, Mänttari M & Manninen V (1994). Relation between QT intervals and heart rates from 40 to 120 beats/min in rest electrocardiograms of men and a simple method to adjust QT interval values. *J Am Coll Cardiol* **23**, 1547–1553.
- Kasemsuppakorn P & Karimi HA (2009). Personalised routing for wheelchair navigation. *J Locat Based Serv* **3**, 24–54.
- Khan MM, Ward R & Ingleby M (2016). Toward Use of Facial Thermal Features in Dynamic Assessment of Affect and Arousal Level. *IEEE Trans Affect Comput PP*, 1–1.
- Kistler A, Mariauzouls C & von Berlepsch K (1998). Fingertip temperature as an indicator for sympathetic responses. *Int J Psychophysiol* **29**, 35–41.
- Kiyatkin EA (2010). Brain Temperature Homeostasis: Physiological Fluctuations and Pathological Shifts. *Front Biosci J Virtual Libr* **15**, 73–92.
- Kjeldstadli K, Tyssen R, Finset A, Hem E, Gude T, Gronvold NT, Ekeberg O & Vaglum P (2006). Life satisfaction and resilience in medical school – a six-year longitudinal, nationwide and comparative study. *BMC Med Educ* **6**, 48.
- Kleinsmith A & Bianchi-Berthouze N (2013). Affective Body Expression Perception and Recognition: A Survey. *IEEE Trans Affect Comput* **4**, 15–33.
- Knoch D, Pascual-Leone A, Meyer K, Treyer V & Fehr E (2006). Diminishing Reciprocal Fairness by Disrupting the Right Prefrontal Cortex. *Science* **314**, 829–832.
- Konrad A, Bellotti V, Crenshaw N, Tucker S, Nelson L, Du H, Pirolli P & Whittaker S (2015). Finding the adaptive sweet spot: Balancing compliance and achievement in automated stress reduction. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pp. 3829–3838. ACM.
- Koob GF, Heinrichs SC, Pich EM, Menzaghi F, Baldwin H, Miczek K & Britton KT (1993). The role of corticotropin-releasing factor in behavioural responses to stress. *Ciba Found Symp* **172**, 277–289; discussion 290-295.

- Krizhevsky A, Sutskever I & Hinton GE (2012). ImageNet Classification with Deep Convolutional Neural Networks. In *Advances in Neural Information Processing Systems 25*, ed. Pereira F, Burges CJC, Bottou L & Weinberger KQ, pp. 1097–1105. Curran Associates, Inc.
- Krzywicki AT, Berntson GG & O’Kane BL (2014). A non-contact technique for measuring eccrine sweat gland activity using passive thermal imaging. *Int J Psychophysiol* **94**, 25–34.
- Kudielka BM, Buske-Kirschbaum A, Hellhammer DH & Kirschbaum C (2004). HPA axis responses to laboratory psychosocial stress in healthy elderly adults, younger adults, and children: impact of age and gender. *Psychoneuroendocrinology* **29**, 83–98.
- Kumar M, Veeraraghavan A & Sabharwal A (2015). DistancePPG: Robust non-contact vital signs monitoring using a camera. *Biomed Opt Express* **6**, 1565–1588.
- Kuraoka K & Nakamura K (2011). The use of nasal skin temperature measurements in studying emotion in macaque monkeys. *Physiol Behav* **102**, 347–355.
- Kurz D (2014). Thermal touch: Thermography-enabled everywhere touch interfaces for mobile augmented reality applications. In *Mixed and Augmented Reality (ISMAR), 2014 IEEE International Symposium on*, pp. 9–16. IEEE.
- Lane ND, Georgiev P & Qendro L (2015). DeepEar: Robust Smartphone Audio Sensing in Unconstrained Acoustic Environments Using Deep Learning. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pp. 283–294. ACM, New York, NY, USA. Available at: <http://doi.acm.org/10.1145/2750858.2804262>.
- Lang P, Bradley M & Cuthbert B (2008). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual*.
- Larson E, Cohn G, Gupta S, Ren X, Harrison B, Fox D & Patel S (2011). HeatWave: Thermal Imaging for Surface User Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 2565–2574. ACM, <http://doi.acm.org/10.1145/1978942.1979317>.
- Laughlin SB, de Ruyter van Steveninck RR & Anderson JC (1998). The metabolic cost of neural information. *Nat Neurosci* **1**, 36–41.
- Lazarus RS (1993). From psychological stress to the emotions: A history of changing outlooks. *Annu Rev Psychol* **44**, 1.
- LeCun Y & Bengio Y (1995). Convolutional Networks for Images, Speech, and Time Series. In *Handbook of Brain Theory and Neural Networks*, ed. Arbib MA, p. 3361. MIT Press.
- LeCun Y, Bengio Y & Hinton G (2015). Deep learning. *Nature* **521**, 436–444.
- Ledda P, Chalmers A, Troscianko T & Seetzen H (2005). Evaluation of Tone Mapping Operators Using a High Dynamic Range Display. In *ACM SIGGRAPH 2005*

Papers, pp. 640–648. ACM, New York, NY, USA. Available at: <http://doi.acm.org/10.1145/1186822.1073242>.

- Lesage F-X, Berjot S & Deschamps F (2012). Clinical stress assessment using a visual analogue scale. *Occup Med* **62**, 600–605.
- Lewis GF, Gatto RG & Porges SW (2011). A novel method for extracting respiration rate and relative tidal volume from infrared thermography. *Psychophysiology* **48**, 877–887.
- Lewis JP (1995). Fast normalized cross-correlation. In *Vision Interface: Canadian Image Processing and Pattern Recognition Society*, pp. 120–123.
- Li C, Narayan RK, Wang P & Hartings JA (2017). Regional temperature and quantitative cerebral blood flow responses to cortical spreading depolarization in the rat. *J Cereb Blood Flow Metab* **37**, 1634–1640.
- Lin M, Chen Q & Yan S (2013). Network In Network. *arXiv preprint arXiv: 13124400*. Available at: <http://arxiv.org/abs/1312.4400>.
- Liu C, Sharan L, Adelson EH & Rosenholtz R (2010). Exploring features in a Bayesian framework for material recognition. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 239–246.
- Liu P, Han S, Meng Z & Tong Y (2014). Facial Expression Recognition via a Boosted Deep Belief Network. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Available at: http://www.cv-foundation.org/openaccess/content_cvpr_2014/html/Liu_Facial_Expression_Recognition_2014_CVPR_paper.html.
- Liu Z & Wang S (2011). Emotion Recognition Using Hidden Markov Models from Facial Temperature Sequence. In *Affective Computing and Intelligent Interaction*, ed. D’Mello S, Graesser A, Schuller B & Martin J-C, Lecture Notes in Computer Science, pp. 240–247. Springer Berlin Heidelberg. Available at: http://link.springer.com/chapter/10.1007/978-3-642-24571-8_26.
- Lloyd JM (2013). *Thermal Imaging Systems*. Springer Science & Business Media.
- Lombard JS (1878a). VI. Experimental researches on the temperature of the head. *Proc R Soc Lond* **27**, 457–465.
- Lombard JS (1878b). II. Experimental researches on the temperature of the head. *Proc R Soc Lond* **27**, 166–177.
- Lucas BD & Kanade T (1981). An Iterative Image Registration Technique with an Application to Stereo Vision. In *Proceedings of the 7th International Joint Conference on Artificial Intelligence - Volume 2*, pp. 674–679. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA. Available at: <http://dl.acm.org/citation.cfm?id=1623264.1623280>.
- Manini B, Cardone D, Ebisch SJH, Bafunno D, Aureli T & Merla A (2013). Mom feels what her child feels: thermal signatures of vicarious autonomic response while

- watching children in a stressful situation. *Front Hum Neurosci*; DOI: 10.3389/fnhum.2013.00299.
- Marazziti D, Di Muro A & Castrogiovanni P (1992). Psychological stress and body temperature changes in humans. *Physiol Behav* **52**, 393–395.
- Martini FH, Timmons MJ & Tallitsch RB (2011). *Human Anatomy*, 7 edition. Pearson, Boston.
- McDuff DJ, Hernandez J, Gontarek S & Picard RW (2016). COGCAM: Contact-free Measurement of Cognitive Stress During Computer Tasks with a Digital Camera. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pp. 4000–4004. ACM, New York, NY, USA. Available at: <http://doi.acm.org/10.1145/2858036.2858247>.
- McEwen BS (2007). Physiology and Neurobiology of Stress and Adaptation: Central Role of the Brain. *Physiol Rev* **87**, 873–904.
- McEwen BS, Bowles NP, Gray JD, Hill MN, Hunter RG, Karatsoreos IN & Nasca C (2015). Mechanisms of stress in the brain. *Nat Neurosci* **18**, 1353–1363.
- McKinley PS, Shapiro PA, Bagiella E, Myers MM, De Meersman RE, Grant I & Sloan RP (2003). Deriving heart period variability from blood pressure waveforms. *J Appl Physiol* **95**, 1431–1438.
- McManus DD, Lee J, Maitas O, Esa N, Pidikiti R, Carlucci A, Harrington J, Mick E & Chon KH (2013). A novel application for the detection of an irregular pulse using an iPhone 4S in patients with atrial fibrillation. *Heart Rhythm* **10**, 315–319.
- Mehrotra A, Tsapeli F, Hendley R & Musolesi M (2017). MyTraces: investigating correlation and causation between users' emotional states and mobile phone interaction. *Proc ACM Interact Mob Wearable Ubiquitous Technol* **1**, 83.
- Mei X & Ling H (2011). Robust Visual Tracking and Vehicle Classification via Sparse Representation. *IEEE Trans Pattern Anal Mach Intell* **33**, 2259–2272.
- Mohan PM, Nagarajan V & Das SR (2016). Stress measurement from wearable photoplethysmographic sensor using heart rate variability data. In *2016 International Conference on Communication and Signal Processing (ICCSP)*, pp. 1141–1144.
- Murthy R & Pavlidis I (2006). Noncontact measurement of breathing function. *IEEE Eng Med Biol Mag* **25**, 57–67.
- Murthy R, Pavlidis I & Tsiamyrtzis P (2004). Touchless monitoring of breathing function. In *26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2004. IEMBS '04*, pp. 1196–1199.
- Naik S & Patel B (2017). Thermal imaging with fuzzy classifier for maturity and size based non-destructive mango (*Mangifera Indica* L.) grading. In *Emerging*

Trends & Innovation in ICT (ICEI), 2017 International Conference on, pp. 15–20. IEEE.

- Nair V & Hinton GE (2010). Rectified linear units improve restricted boltzmann machines. In *Proceedings of the 27th international conference on machine learning (ICML-10)*, pp. 807–814. Available at: http://machinelearning.wustl.edu/mlpapers/paper_files/icml2010_NairH10.pdf.
- Nakayama K, Goto S, Kuraoka K & Nakamura K (2005). Decrease in nasal temperature of rhesus monkeys (*Macaca mulatta*) in negative emotional state. *Physiol Behav* **84**, 783–790.
- Nash JM & Theborge RW (2006). Understanding Psychological Stress, Its Biological Processes, and Impact on Primary Headache. *Headache J Head Face Pain* **46**, 1377–1386.
- Nhan BR & Chau T (2010). Classifying Affective States Using Thermal Infrared Imaging of the Human Face. *IEEE Trans Biomed Eng* **57**, 979–987.
- Or CKL & Duffy VG (2007). Development of a facial skin temperature-based methodology for non-intrusive mental workload measurement. *Occup Ergon* **7**, 83–94.
- Pagani M, Lombardi F, Guzzetti S, Rimoldi O, Furlan R, Pizzinelli P, Sandrone G, Malfatto G, Dell’Orto S & Piccaluga E (1986). Power spectral analysis of heart rate and arterial pressure variabilities as a marker of sympatho-vagal interaction in man and conscious dog. *Circ Res* **59**, 178–193.
- Pariante CM & Lightman SL (2008). The HPA axis in major depression: classical theories and new developments. *Trends Neurosci* **31**, 464–468.
- Patton KT (2015). *Anatomy and Physiology*. Elsevier Health Sciences.
- Pavlidis I, Eberhardt NL & Levine JA (2002). Human behaviour: Seeing through the face of deception. *Nature* **415**, 35–35.
- Pavlidis I, Levine J & Baukol P (2001). Thermal image analysis for anxiety detection. In *2001 International Conference on Image Processing, 2001. Proceedings*, pp. 315–318 vol.2.
- Pavlidis I, Symosek P, Fritz B, Bazakos M & Papanikolopoulos N (2000). Automatic detection of vehicle occupants: the imaging problem and its solution. *Mach Vis Appl* **11**, 313–320.
- Pavlidis I, Tsiamyrtzis P, Shastri D, Wesley A, Zhou Y, Lindner P, Buddharaju P, Joseph R, Mandapati A, Dunkin B & Bass B (2012). Fast by Nature - How Stress Patterns Define Human Experience and Performance in Dexterous Tasks. *Sci Rep*; DOI: 10.1038/srep00305.
- Pearson J (1786). Observations and Queries on Animal Heat. *Lond Med J* **7**, 169–180.

- Pereira CB, Yu X, Czaplik M, Rossaint R, Blazek V & Leonhardt S (2015). Remote monitoring of breathing dynamics using infrared thermography. *Biomed Opt Express* **6**, 4378–4394.
- Pergola PE, Kellogg DL, Johnson JM, Kosiba WA & Solomon DE (1993). Role of sympathetic nerves in the vascular effects of local temperature in human forearm skin. *Am J Physiol-Heart Circ Physiol* **265**, H785–H792.
- Plews DJ, Scott B, Altini M, Wood M, Kilding AE & Laursen PB (2017). Comparison of Heart-Rate-Variability Recording With Smartphone Photoplethysmography, Polar H7 Chest Strap, and Electrocardiography. *Int J Sports Physiol Perform* **12**, 1324–1328.
- Poh MZ, McDuff DJ & Picard RW (2011). Advancements in Noncontact, Multiparameter Physiological Measurements Using a Webcam. *IEEE Trans Biomed Eng* **58**, 7–11.
- Pollina DA, Dollins AB, Senter SM, Brown TE, Pavlidis I, Levine JA & Ryan AH (2006). Facial Skin Surface Temperature Changes During a “Concealed Information” Test. *Ann Biomed Eng* **34**, 1182–1189.
- Pruessner JC, Champagne F, Meaney MJ & Dagher A (2004). Dopamine Release in Response to a Psychological Stress in Humans and Its Relationship to Early Life Maternal Care: A Positron Emission Tomography Study Using [11C]Raclopride. *J Neurosci* **24**, 2825–2831.
- Pruessner JC, Dedovic K, Khalili-Mahani N, Engert V, Pruessner M, Buss C, Renwick R, Dagher A, Meaney MJ & Lupien S (2008). Deactivation of the Limbic System During Acute Psychosocial Stress: Evidence from Positron Emission Tomography and Functional Magnetic Resonance Imaging Studies. *Biol Psychiatry* **63**, 234–240.
- Ptakauskaitė N, Cox AL & Berthouze N (2018). Knowing What You’re Doing or Knowing what to do: How Stress Management Apps Support Reflection and Behaviour Change. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, p. LBW599. ACM.
- Puri C, Olson L, Pavlidis I, Levine J & Starren J (2005). StressCam: Non-contact Measurement of Users’ Emotional States Through Thermal Imaging. In *CHI ’05 Extended Abstracts on Human Factors in Computing Systems*, pp. 1725–1728. ACM, <http://doi.acm.org/10.1145/1056808.1057007>.
- Ren S, He K, Girshick R & Sun J (2015). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. In *Advances in Neural Information Processing Systems* 28, ed. Cortes C, Lawrence ND, Lee DD, Sugiyama M & Garnett R, pp. 91–99. Curran Associates, Inc.
- Ridler T & Calvard S (1978). Picture Thresholding Using an Iterative Selection Method. *IEEE Trans Syst Man Cybern* **8**, 630–632.

- Ring EFJ & Ammer K (2012). Infrared thermal imaging in medicine. *Physiol Meas* **33**, R33.
- Ringer S (1873). *On the Temperature of the Body as a Means of Diagnosis and Prognosis in Phthisis*. Lewis.
- Ropers D, Baum U, Pohle K, Anders K, Ulzheimer S, Ohnesorge B, Schlundt C, Bautz W, Daniel WG & Achenbach S (2003). Detection of Coronary Artery Stenoses With Thin-Slice Multi-Detector Row Spiral Computed Tomography and Multiplanar Reconstruction. *Circulation* **107**, 664–666.
- Rosenbaum D, Hilsendegen P, Thomas M, Haeussinger FB, Metzger FG, Nuerk H-C, Fallgatter AJ, Nieratschker V & Ehlis A-C (2018). Cortical hemodynamic changes during the Trier Social Stress Test: An fNIRS study. *NeuroImage* **171**, 107–115.
- Roychoudhuri C, Kracklauer AF & Creath K (2008). *The nature of light: what is a photon?* CRC Press.
- Sahami Shirazi A, Abdelrahman Y, Henze N, Schneegass S, Khalilbeigi M & Schmidt A (2014). Exploiting Thermal Reflection for Interactive Systems. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems*, pp. 3483–3492. ACM, New York, NY, USA. Available at: <http://doi.acm.org/10.1145/2556288.2557208>.
- Salahuddin L & Kim D (2006). Detection of Acute Stress by Heart Rate Variability Using a Prototype Mobile ECG Sensor. In *Proceedings of the 2006 International Conference on Hybrid Information Technology - Volume 02*, pp. 453–459. IEEE Computer Society, Washington, DC, USA. Available at: <http://dx.doi.org/10.1109/ICHIT.2006.117>.
- Salazar-López E, Domínguez E, Juárez Ramos V, de la Fuente J, Meins A, Iborra O, Gálvez G, Rodríguez-Artacho MA & Gómez-Milán E (2015). The mental and subjective skin: Emotion, empathy, feelings and thermography. *Conscious Cogn* **34**, 149–162.
- Salmivalli C, Kaukiainen A, Kaistaniemi L & Lagerspetz KMJ (1999). Self-Evaluated Self-Esteem, Peer-Evaluated Self-Esteem, and Defensive Egotism as Predictors of Adolescents' Participation in Bullying Situations. *Pers Soc Psychol Bull* **25**, 1268–1278.
- Sano A & Picard RW (2013). Stress Recognition Using Wearable Sensors and Mobile Phones. In *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*, pp. 671–676.
- Sato M, Yoshida S, Olwal A, Shi B, Hiyama A, Tanikawa T, Hirose M & Raskar R (2015). SpecTrans: Versatile Material Classification for Interaction with Textureless, Specular and Transparent Surfaces. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pp. 2191–2200. ACM, <http://doi.acm.org/10.1145/2702123.2702169>.

- Schäfer A & Vagedes J (2013). How accurate is pulse rate variability as an estimate of heart rate variability? A review on studies comparing photoplethysmographic technology with an electrocardiogram. *Int J Cardiol* **166**, 15–29.
- Schmidt A (2002). *Ubiquitous computing - computing in context* (Ph.D. thesis). Lancaster University. Available at: <http://eprints.lancs.ac.uk/12221/>.
- Schmidt A, Beigl M & Gellersen H-W (1999). There is more to context than location. *Comput Graph* **23**, 893–901.
- Seematter G, Dirlwanger M, Rey V, Schneiter P & Tappy L (2002). Metabolic Effects of Mental Stress during Over- and Underfeeding in Healthy Women. *Obes Res* **10**, 49–55.
- Selye H (1980). *Selye's guide to stress research*. Van Nostrand Reinhold.
- Setz C, Arnrich B, Schumm J, La Marca R, Tröster G & Ehlert U (2010). Discriminating stress from cognitive load using a wearable EDA device. *IEEE Trans Inf Technol Biomed* **14**, 410–417.
- Shaffer F & Ginsberg JP (2017). An Overview of Heart Rate Variability Metrics and Norms. *Front Public Health*; DOI: 10.3389/fpubh.2017.00258.
- Shaffer F, McCraty R & Zerr CL (2014). A healthy heart is not a metronome: an integrative review of the heart's anatomy and heart rate variability. *Front Psychol*; DOI: 10.3389/fpsyg.2014.01040.
- Shannon CE (1948). A mathematical theory of communication. *Bell Syst Tech J* **27**, 379–423.
- Shastri D, Merla A, Tsiamyrtzis P & Pavlidis* I (2009). Imaging Facial Signs of Neurophysiological Responses. *IEEE Trans Biomed Eng* **56**, 477–484.
- Shastri D, Papadakis M, Tsiamyrtzis P, Bass B & Pavlidis I (2012). Perinasal Imaging of Physiological Stress and Its Affective Potential. *IEEE Trans Affect Comput* **3**, 366–378.
- Shi J & Tomasi C (1994). Good features to track. In *1994 Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 593–600.
- Shin H (2016). Ambient temperature effect on pulse rate variability as an alternative to heart rate variability in young adult. *J Clin Monit Comput* **30**, 939–948.
- Smets EMA, Garssen B, Bonke B & De Haes JCJM (1995). The multidimensional Fatigue Inventory (MFI) psychometric qualities of an instrument to assess fatigue. *J Psychosom Res* **39**, 315–325.
- Sonka M, Hlavac V & Boyle R (2014). *Image Processing, Analysis, and Machine Vision*. Cengage Learning.

- Soufer R, Bremner JD, Arrighi JA, Cohen I, Zaret BL, Burg MM & Goldman-Rakic P (1998). Cerebral cortical hyperactivation in response to mental stress in patients with coronary artery disease. *Proc Natl Acad Sci* **95**, 6454–6459.
- Spalding SJ, Kwok CK, Boudreau R, Enama J, Lunich J, Huber D, Denes L & Hirsch R (2008). Three-dimensional and thermal surface imaging produces reliable measures of joint shape and temperature: a potential tool for quantifying arthritis. *Arthritis Res Ther* **10**, R10.
- Srivastava N, Hinton GE, Krizhevsky A, Sutskever I & Salakhutdinov R (2014). Dropout: a simple way to prevent neural networks from overfitting. *J Mach Learn Res* **15**, 1929–1958.
- Steketee J (1973). Spectral emissivity of skin and pericardium. *Phys Med Biol* **18**, 686–694.
- Stroop JR (1935). Studies of interference in serial verbal reactions. *J Exp Psychol* **18**, 643–662.
- Sun N, Garbey M, Merla A & Pavlidis I (2005). Imaging the cardiovascular pulse. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, pp. 416–421 vol. 2.
- Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V & Rabinovich A (2015). Going Deeper With Convolutions. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Available at: http://www.cv-foundation.org/openaccess/content_cvpr_2015/html/Szegedy_Going_Deeper_With_2015_CVPR_paper.html [Accessed March 24, 2017].
- Taylor SF, Kornblum S, Lauber EJ, Minoshima S & Koeppe RA (1997). Isolation of Specific Interference Processing in the Stroop Task: PET Activation Studies. *NeuroImage* **6**, 81–92.
- Tetko IV, Livingstone DJ & Luik AI (1995). Neural network studies. 1. Comparison of overfitting and overtraining. *J Chem Inf Comput Sci* **35**, 826–833.
- Trigeorgis G, Ringeval F, Brueckner R, Marchi E, Nicolaou MA, Schuller B & Zafeiriou S (2016). Adieu features? End-to-end speech emotion recognition using a deep convolutional recurrent network. In *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 5200–5204.
- Tsiamyrtzis P, Dowdall J, Shastri D, Pavlidis IT, Frank MG & Ekman P (2007). Imaging Facial Physiology for the Detection of Deceit. *Int J Comput Vis* **71**, 197–214.
- Tsuji H, Venditti FJ, Manders ES, Evans JC, Larson MG, Feldman CL & Levy D (1996). Determinants of heart rate variability. *J Am Coll Cardiol* **28**, 1539–1546.
- Tukey J (1977). *Exploratory Data Analysis*. Pearson.
- Ursin H & Eriksen H (2007). Cognitive Activation Theory of Stress, Sensitization, and Common Health Complaints. *Ann N Y Acad Sci* **1113**, 304–310.

- Vedaldi A & Lenc K (2015). MatConvNet: Convolutional Neural Networks for MATLAB. In *Proceedings of the 23rd ACM International Conference on Multimedia*, pp. 689–692. ACM, New York, NY, USA. Available at: <http://doi.acm.org/10.1145/2733373.2807412>.
- Veltman JA, Vos WK & TNO Defensie en Veiligheid (2005). Facial temperature as a measure of mental workload. In *International Symposium on Aviation Psychology*. Oklahoma City.
- Verkrusye W, Svaasand LO & Nelson JS (2008). Remote plethysmographic imaging using ambient light. *Opt Express* **16**, 21434–21445.
- Vizer LM, Zhou L & Sears A (2009). Automated stress detection using keystroke and linguistic features: An exploratory study. *Int J Hum-Comput Stud* **67**, 870–886.
- Vollmer M & Klaus-Peter M (2017). *Infrared thermal imaging: fundamentals, research and applications*. John Wiley & Sons.
- Wallin BG (1981). Sympathetic Nerve Activity Underlying Electrodermal and Cardiovascular Reactions in Man. *Psychophysiology* **18**, 470–476.
- Wang H, Wang B, Normoyle KP, Jackson K, Spitler K, Sharrock MF, Miller CM, Best C, Llano D & Du R (2014a). Brain temperature and its fundamental properties: a review for clinical neuroscientists. *Syst Biol* **8**, 307.
- Wang J, Rao H, Wetmore GS, Furlan PM, Korczykowski M, Dinges DF & Detre JA (2005). Perfusion functional MRI reveals cerebral blood flow pattern under psychological stress. *Proc Natl Acad Sci U S A* **102**, 17804–17809.
- Wang J-S, Lin C-W & Yang Y-TC (2013). A k-nearest-neighbor classifier with heart rate variability feature-based transformation algorithm for driving stress recognition. *Neurocomputing* **116**, 136–143.
- Wang S, He M, Gao Z, He S & Ji Q (2014b). Emotion recognition from thermal infrared images using deep Boltzmann machine. *Front Comput Sci* **8**, 609–618.
- Wang S, Liu Z, Lv S, Lv Y, Wu G, Peng P, Chen F & Wang X (2010). A Natural Visible and Infrared Facial Expression Database for Expression Recognition and Emotion Inference. *IEEE Trans Multimed* **12**, 682–691.
- Wang T-C, Zhu J-Y, Hiroaki E, Chandraker M, Efros AA & Ramamoorthi R (2016). A 4D Light-Field Dataset and CNN Architectures for Material Recognition. In *Computer Vision – ECCV 2016*, pp. 121–138. Springer, Cham. Available at: https://link.springer.com/chapter/10.1007/978-3-319-46487-9_8.
- Wasner G, Schattschneider J & Baron R (2002). Skin temperature side differences – a diagnostic tool for CRPS? *Pain* **98**, 19–26.
- Weinberger DR, Berman K & Zec RF (1986). Physiologic dysfunction of dorsolateral prefrontal cortex in schizophrenia: I. regional cerebral blood flow evidence. *Arch Gen Psychiatry* **43**, 114–124.

- Wesley A, Buddharaju P, Pienta R & Pavlidis I (2012). A Comparative Analysis of Thermal and Visual Modalities for Automated Facial Expression Recognition. In *Advances in Visual Computing*, Lecture Notes in Computer Science, pp. 51–60. Springer Berlin Heidelberg.
- West JB (2012). *Respiratory Physiology: The Essentials*. Lippincott Williams & Wilkins.
- White RD & Flaker G (2017). Smartphone-based Arrhythmia Detection: Should we encourage patients to use the ECG in their pocket? *J Atr Fibrillation*.
- Williamon A, Aufegger L, Wasley D, Looney D & Mandic DP (2013). Complexity of physiological responses decreases in high-stress musical performance. *J R Soc Interface* **10**, 20130719.
- Xu S, Sun L & Rohde GK (2014). Robust efficient estimation of heart rate pulse from video. *Biomed Opt Express* **5**, 1124–1135.
- Yannakakis GN, Cowie R & Busso C (2017). The ordinal nature of emotions. In *2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII)*, pp. 248–255.
- Yarkoni T, Poldrack RA, Nichols TE, Van Essen DC & Wager TD (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nat Methods* **8**, 665.
- Yarmolenko PS, Moon EJ, Landon C, Manzoor A, Hochman DW, Viglianti BL & Dewhirst MW (2011). Thresholds for thermal damage to normal tissues: An update. *Int J Hyperth Off J Eur Soc Hyperthermic Oncol North Am Hyperth Group* **27**, 320–343.
- Yeo H-S, Flamich G, Schrempf P, Harris-Birtill D & Quigley A (2016). RadarCat: Radar Categorization for Input & Interaction. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, pp. 833–841. ACM, <http://doi.acm.org/10.1145/2984511.2984515>.
- Yeo H-S, Lee J, Bianchi A, Harris-Birtill D & Quigley A (2017). SpeCam: Sensing Surface Color and Material with the Front-facing Camera of a Mobile Device. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services*, pp. 25:1–25:9. ACM, New York, NY, USA. Available at: <http://doi.acm.org/10.1145/3098279.3098541>.
- Yerkes RM & Dodson JD (1908). The relation of strength of stimulus to rapidity of habit-formation. *J Comp Neurol* **18**, 459–482.
- Yu B, Funk M, Hu J, Wang Q & Feijs L (2018). Biofeedback for Everyday Stress Management: A Systematic Review. *Front ICT*; DOI: 10.3389/fict.2018.00023.
- Yu Z & Zhang C (2015). Image Based Static Facial Expression Recognition with Multiple Deep Network Learning. In *Proceedings of the 2015 ACM on*

International Conference on Multimodal Interaction, pp. 435–442. ACM, New York, NY, USA. Available at: <http://doi.acm.org/10.1145/2818346.2830595>.

Zhou Y, Tsiamyrtzis P & Pavlidis IT (2009). Tissue Tracking in Thermo-physiological Imagery through Spatio-temporal Smoothing. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2009*, Lecture Notes in Computer Science, pp. 1092–1099. Springer Berlin Heidelberg. Available at: http://link.springer.com/chapter/10.1007/978-3-642-04271-3_132.

Zhu B, Hedman A, Feng S, Li H & Osika W (2017). Designing, Prototyping and Evaluating Digital Mindfulness Applications: A Case Study of Mindful Breathing for Stress Reduction. *J Med Internet Res* **19**, e197.

Ziegler LH & Cash PT (1938). A study of the influence of emotions and affects on the surface temperature of the human body. *Am J Psychiatry* **95**, 677–696.

Appendices

Appendix A. (Extra Study) Skeletal Muscle Activation as a New Thermal Signature from Underexplored ROIs: Understanding of Thermal Signatures of Physical Stress during Anaerobic Exercise

Together with mental distress, physical stress is an important physiological phenomenon and one of the facets of stress (Ursin & Eriksen, 2007). Even though this thesis mainly focuses on a psychological facet of stress, the ability to detect physical stress would provide a variety of benefits. In sports for example, people are likely to overdo exercises without respect to fatigue or physical stress, probably leading to injury. Providing sensory feedback concerning physical overload cues could help in such situations and could be used in rehabilitation and physical therapy. Hence, we review the literature of muscular physiology of heat production. We then conduct a pilot study, and report results, to explore temperature variations in healthy adults who attend anaerobic exercise sessions to induce repeated muscle contractions leading to physical stress.

A.1. Background: Physical Stress and Muscle Temperature

Although the scope of this thesis is limited to *psychological stress*, this section provides reviews of physiology of muscle temperature in relation to *physical stress* and then discusses further insights.

Given that it is difficult to attain perfect efficiency in the catabolic processes of cells in human beings, some of the energy is lost which causes heat transformation. Skeletal muscle, the voluntary muscle tissue in the human body which can be controlled by our will, strongly contributes to the production of a high amount of heat, helping to maintain our regular bodily temperature (Patton, 2015; Hall & Guyton, 2015).

The relationship between muscle stress and heat production can be explained by the balance between energy demand and supply. Bringing forward the term of biochemical reactions, (Barclay *et al.*, 1995) assessed the energy supply and demand by measuring muscle heat production, calling phosphocreatine (PC_r) breakdown and its re-synthesis as the ‘initial’ and ‘recovery’ process, respectively. The results of their experiment indicate that in non-fatigued muscle adenosine triphosphate (ATP), which is a key energy source to the organism activation, this would be rapidly resynthesized by a series of biochemical reaction, an imbalance between the demand for the PC_r , and its supply which causes muscle fatigue.

Each contraction of a muscle instantly produces the ‘initial’ heat on the muscle, and the ‘recovery’ heat is followed at a relatively slow rate between contractions. The ratio between the initial and recovery heats (‘recovery’ per ‘initial’ heat) increases steeply at early contractions, however it comes across a stable status in successive contraction cycles as if it encounters the steady-state after the transient response in the electric circuit or mechanical structure. The authors identified the mean value of the ratio over the last ten cycles as proportional to a decline of force produced by the muscle in accordance with contraction sequences. The insufficient amount of resynthesized PC_r causes the feeling of fatigue and presents an amount of heat change. Although they targeted mouse skeletal muscles, the findings could be applied to human beings. It is expected to possibly observe the specific tendency of heat production in human skeletal muscle in accordance with mismatch between energy demand and supply process caused by the excessive muscle contraction arousing physical and general fatigue.

In particular, muscle contraction plays a central role in the production of heat on the skeletal tissues and needs an energy source called ATP. The biochemical reaction referred to as creatine phosphokinase reaction can be expressed as



During the reaction, an amount of heat is produced. In terms of metabolic expense, neuronal activity requires a metabolically high cost. According to Laughlin *et al.* (1998), approximately 10^4 ATP molecules are used to transmit a slight amount of information at a chemical synapse, and almost 10^6 - 10^7 ATPs are needed to recharge a neuron for a single action potential. Certainly, this is followed by the generation of a considerable amount of heat.

Prolonged and strong contraction of the skeletal muscle leads to physical fatigue. According to Smets *et al.* (1995), a general fatigue, which is a subjective feeling of tiredness and physical fatigue, is related to muscular distress. This is highly inter-correlated regardless of

different individual occupations or health conditions. The fatigue is accompanied by a decrease in performance of both psychological and behavioural activities and results in a failure to sustain the needed behaviour (Smets *et al.*, 1995). Among several factors contributing to the feeling of fatigue, such as the central nervous system activity, sarcolemma and calcium, the by-products related to ATP present their main contribution towards fatigue during muscle activity.

The process of physical fatigue can be explained through a mechanism relevant to ATP synthesis. Firstly, muscular activity continually requires energy sources, and steadily lessens the ability of the muscular tissues to reproduce ATP at a quicker pace. This leads to a shortage of ATP, making myosin heads incapable of producing the necessary force for prolonged muscle contraction and causing physiological muscle fatigue. In addition, blood flow through a muscle is repeatedly interrupted during the prolonged muscle contractions, causing the loss of nutrient supply, as well as loss of oxygen or glucose in muscle fibres. This in turn leads to low levels of ATP due to the loss of nutrients disrupting re-synthesis of ATP, resulting in making people feel fatigue (Patton, 2015; Hall & Guyton, 2015).

As for the relationship between the physical muscle fatigue and psychological fatigue, Patton (2015) documented that while performing mild physical activities, a person hardly experiences “*complete*” physical fatigue (or stress) which makes him/her unable to contract muscles anymore and what the person feels is highly likely to be psychological fatigue. Instead, psychological fatigue (this term is defined as “*a subjective sensation characterised by lack of motivation and of alertness*”), according to Sharpe (1991) and Smets *et al.*, (1995), provokes mental tiredness which makes a person discontinuing a physical movement (Patton, 2015). Therefore, it is expected that a pattern of heat production during a muscular activity could provide thermal signatures of human physical stress, and particular skeletal muscles activated in physical activities could be valid ROIs as shown in Figure A.1.

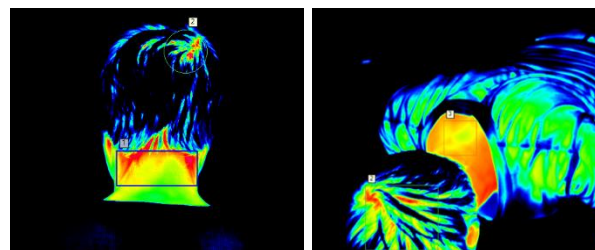


Figure A.1. Superficial muscle dissection of the neck and back.

Empirically, Bartuzi *et al.* (2012) analysed the effect of physical stress on the temperature variance of muscles using thermography and the superficial Electromyography (EMG) signal. The EMG has been broadly employed in measurements of muscular activations and is one of key techniques to assess muscle fatigue (physical stress). The authors conducted the study with 10 male participants performing 5%, 15%, and 30% of maximum voluntary contraction (MVC) lasting 5 minutes with biceps brachii of their right upper limb. As a result, they documented the root mean square of the EMG scores as showing a statistically positive correlation with thermal values on the biceps brachii muscle, and the mean power frequency and median frequency of the EMG were negatively correlated with the temperature.

In addition, they identified the higher the level of muscle contraction a participant performs, the higher the increase of temperature can be observed. However, the experiment was limited to a one-off contraction, not the repeated contractions, so further exploration should be followed in order to understand detailed patterns of thermal signatures of fatigue and to bring the thermography-based measurement into our daily lives.

A.2. Experimental Protocol and Dataset

In this study, a push up, one example of numerous popular anaerobic exercises, is employed to find thermal signatures in association with fatigue. In particular, the pilot study concentrates on thermal cues of the upper trapezius muscle which is actively contracted during the exercise (Calatayud *et al.*, 2014).

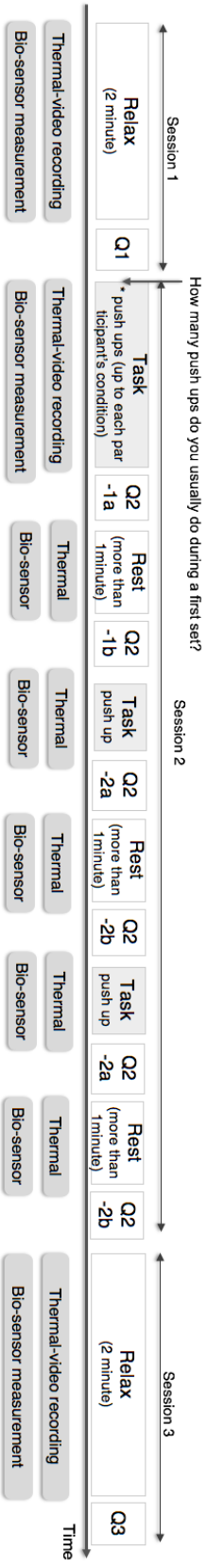
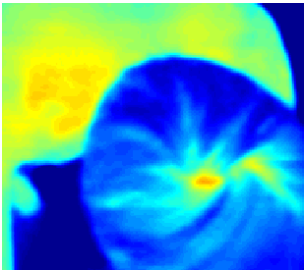
In the pilot study, 4 healthy adults (1 female) (aged 25-31 years, $M=28.75$, $SD=2.63$) were recruited from the University College London and non-research community. Each subject was given the information sheet and the informed consent form prior to data acquisition. The experiment was conducted in a quiet lab room with no distractions and took approximately half an hour per participant.

The task was split into three segments: a rest period (2-minutes), 3 times repeated exercise including a set of push-ups as well as a rest period, and a post-rest (3-minutes). The segments are shown in Figure A.1. During each session all participants were asked to fill out a questionnaire which consists of demographics and question-sets which were recorded by a thermal camera (see Figure 1.2b). Specifically, the left-upper trapezius muscle undergoing contraction and relaxation

along a set of push-ups were centrally focused. Before the start of the exercise, each subject was asked “how many push ups do you usually do during each set” in order to provide the maximum number of push-up repetitions. In addition, participants were required to take off their shirts (male) or wear light sportswear (female) and perform the push-ups in the same pose and location, as well as making sure their physical conditions were fully recovered during the rest period before starting the next segment.

After the end of each rest period, every subject was asked to answer three questions based on a ten score Visual Analogue Scale (VAS), which is a straight line allowing participants to answer on an analogue basis, instead of using the Likert scale, in order to avoid its non-parametric properties (Bijur *et al.*, 2001). The three questions were: 1) *Do you feel mentally fatigued or tiredness? (Psychological fatigue is a subjective sensation characterized by lack of motivation and of alertness) (ranging from 0, no fatigue, to 10, total exhaustion)*, 2) *Do you feel physically tired or fatigued? (ranging from 0, no fatigue, to 10, total exhaustion)*, and 3) *Do you feel relaxed? (ranging from 0, not at all, to 10, very much)*.

After the end of each set of the exercise, every subject was required to answer six questions based on the VAS which were: 1) Did you feel mentally fatigued or tiredness? (Psychological fatigue is a subjective sensation characterized by lack of motivation and of alertness) (ranging from 0, no fatigue, to 10, total exhaustion), 2) Did you feel physically tired or fatigued? (ranging from 0, no fatigue, to 10, total exhaustion), 3) Difficulty of the task? (ranging from 0, low, to 10, high), 4) How much do you feel you engaged in the task? (ranging from 0, low, to 10, high), 5) How much do you think you performed in the task? (ranging from 0, low, to 10, high), and 6) Did you feel pain in your muscles? If so, where? and how much? (0-10). The last question was designed to investigate interaction effects between fatigue and pain conditions for special use.



* Rest period: Make sure your physical condition is fully recovered

Figure A.2. The task-flow of the study and an example of thermal images collected during the experiment.

A.3. Results

Figure A.3 shows a series of thermal images which were subsequently collected while a subject (P2) was performing a trial of push-ups. The upper trapezius muscle activation can be thermally observed as depicted in Figure A.3b. The activation could be possibly quantified by using statistical features such as the integral-ratio based on thresholding and skewness, and information theoretic features such as entropy. Although a significance test should be done as part of future work, different values of the features were gathered in accordance with thermal activation patterns on the upper trapezius muscle as shown in Figure A.4. Thermograms on the upper trapezius muscles collected during a push-up session were presented in Figure A.3e. While the maximum temperature on the ROI is not able to provide the full information about two-dimensional thermal distribution in relation to muscle activation (Figure A.4a), the integral ratio, skewness, and entropy reflect the dimensionality and could be selected as main features indicating the skeletal muscle activation. In other words, these features can be used in automatic recognition of fatigue in case there are significant differences between each feature.

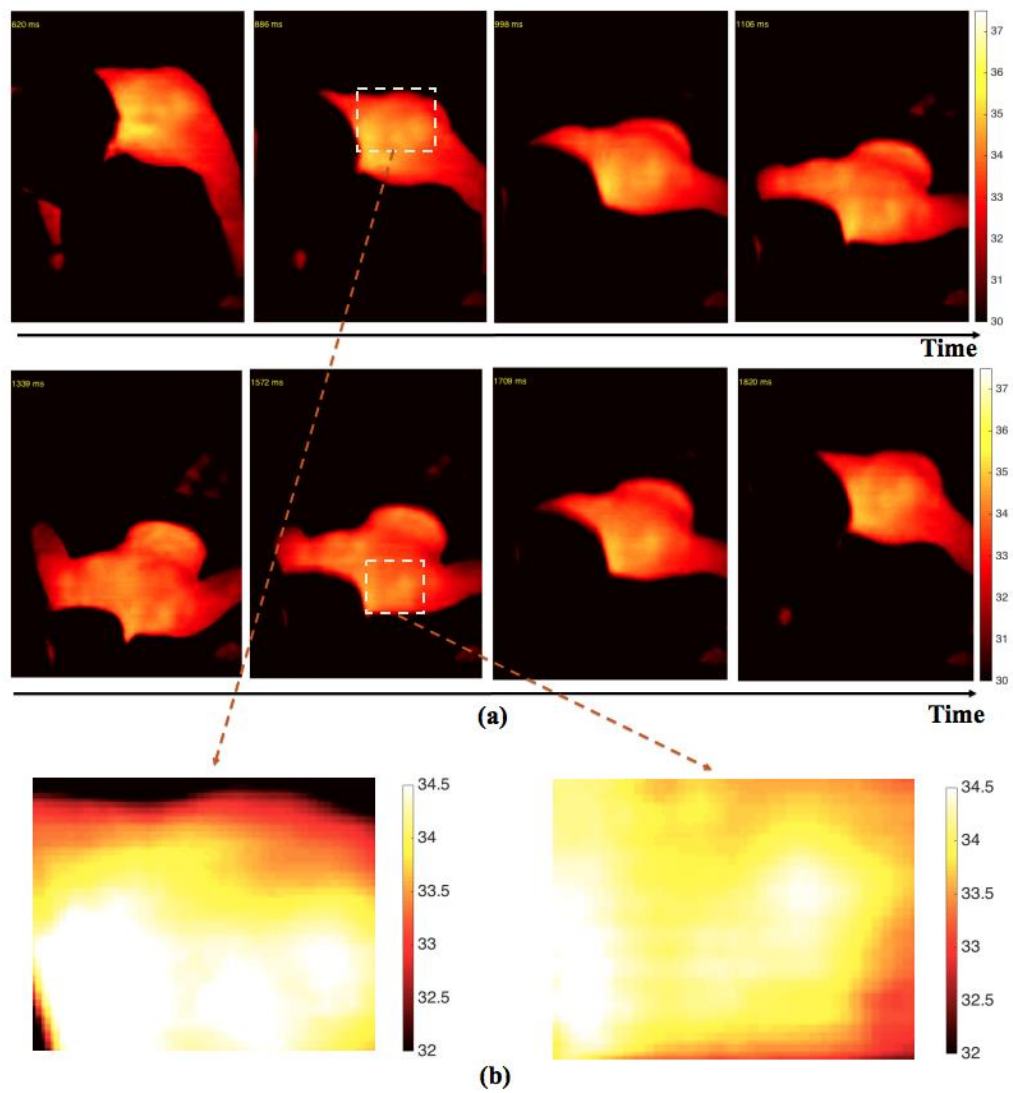


Figure A.3. (a) thermograms collected during the push-up sessions (Participant 2), (b) ROI: thermal patterns on the upper trapezius muscle in a narrow range of temperature – just before ‘push’ (left), ‘up’ (right).

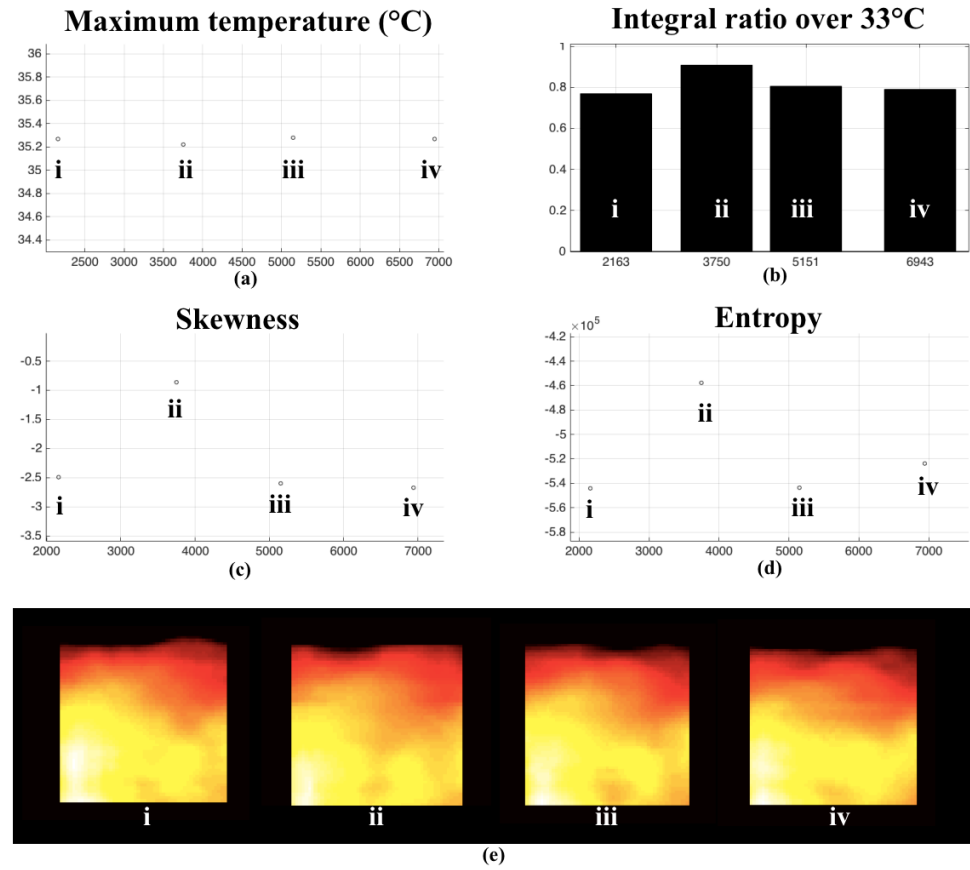


Figure A.4. (a) maximum temperature, (b) integral ratio based on thresholding, (c) skewness, (d) entropy on (e) a specific ROI (upper trapezius muscle) (Participant 2).

In consideration of findings from experiments which investigated mice' heat production during muscle contraction by Barclay *et al.*, 1995, initial heat and recovery heat amounts were supposed to be measured from different moments, just before 'up' (Figure A.3b, right), and 'push' (Figure A.4b, left), respectively. However, the concepts (i.e. initial-recovery) require a relatively long pause between each contraction to wait for the heat production. In this experiment, participants executed the exercise at natural speed, so the maximum thermal signals from the moments shown in Figure A.3b. (left) were instead collected manually to observe tendencies on heat production in relation to general fatigue. Table A.1 summarises the overall results. According to this result, heat appears to be roughly produced from the activated muscles along muscle contraction which was gradually saturated and started to decrease around the moment participants felt most fatigued. The correlation coefficient between the differential subjective rating on fatigue and the differential maximum temperature produced on the upper trapezius was -0.341.

Table A.1. Comparison between differential subjective ratings based on the VAS scale and difference of maximum temperature on the upper trapezius muscle during push-ups and rest periods.

<i>Participants</i>	<i>Score</i>	<i>Base-line</i>	<i>Push-ups (#1)</i>	<i>Rest (#1)</i>	<i>Push-ups (#2)</i>	<i>Rest (#2)</i>	<i>Push-ups (#3)</i>	<i>Rest (#3)</i>	<i>Relax</i>
P1	<i>Rating on general fatigue (VAS scale)</i>	2.2	3.5	2	3.8	2.2	3.6	3.2	4
	<i>Differential Rating (ΔVAS)</i>		1.3	-1.5	1.8	-1.6	1.4	-0.4	0.8
	<i>Differential maximum temperature</i>		-0.3	0.3	-0.65	-0.13	-0.34	-0.05	-0.36
P2	<i>Rating on general fatigue (VAS scale)</i>	5	5.1	4	5	5	6	5	6
	<i>Differential Rating (ΔVAS)</i>		0.1	-1.1	1	0	1	-1	1
	<i>Differential maximum temperature</i>		0.08	-0.31	0.98	0.1	-0.25	0.55	-0.26
P3	<i>Rating on general fatigue (VAS scale)</i>	4.6	5.9	3	7	4	9	7	6.3
	<i>Differential Rating (ΔVAS)</i>		1.3	-2.9	4	-3	5	-2	-0.7
	<i>Differential maximum temperature</i>		-0.23	-0.17	-0.26	0.16	-0.1	-0.39	0.39
P4	<i>Rating on general fatigue (VAS scale)</i>	4	6.6	5	6.4	4.7	8.4	7.4	5.7
	<i>Differential Rating (ΔVAS)</i>		2.6	-1.6	1.4	-1.7	3.7	-1	-1.7
	<i>Differential maximum temperature</i>		-0.5	0.05	-0.43	0.26	-0.24	0.08	0.01

Interestingly, it is noteworthy that there were specific patterns of heat variation produced during the rest periods after the completion of each set of push ups. As can be seen in Figure A.5. (maximum temperature on the upper trapezius muscle), a higher amount of heat was steadily created even while the participant was taking a break between push-ups sets. The subject responded with feeling a high level of fatigue on the VAS questionnaire after finishing the last set of push-ups, and a drop in the maximum temperature on the muscle was found during the last rest period. In parallel, the muscle deactivation could also be found, as illustrated in Figure A.6.

However, more experiments should be done to identify the thermal signatures of fatigue during relaxation.

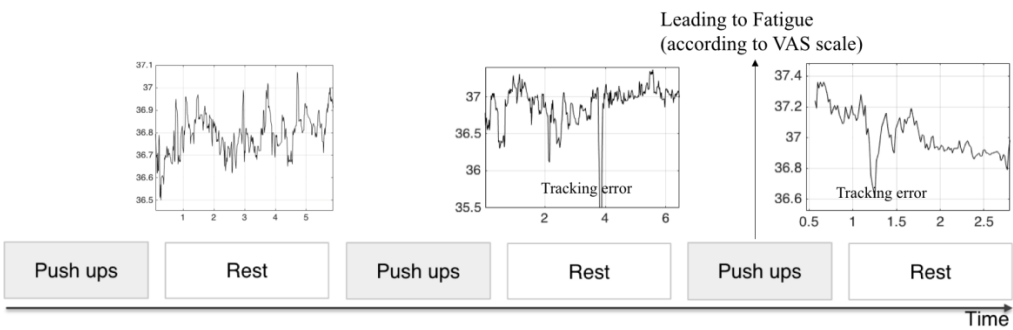


Figure A.5. Maximum temperature sequences on the upper trapezius muscle during the rest periods after the completion of each set of push ups (Participant 2).

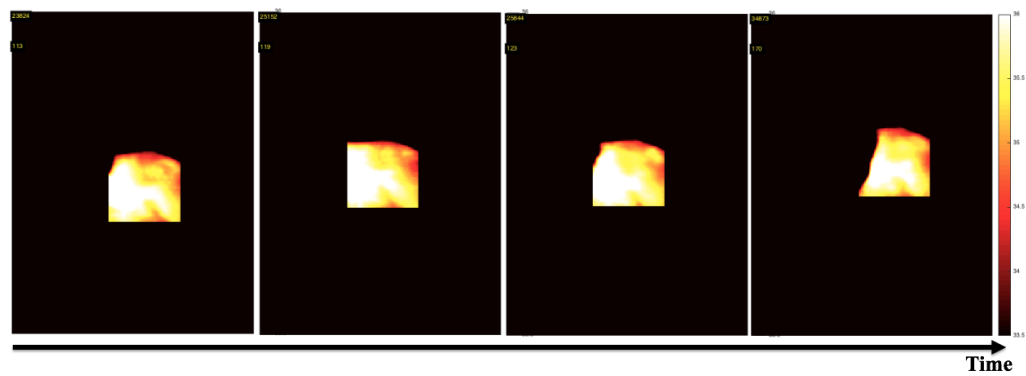


Figure A.6. Upper trapezius muscle activations during the rest period (participant 2).

Lastly, there was a concerning point about the physical setting of the thermal camera. Because the recording position was fixed, distance to the muscle and viewing angles were severely changed along with each moment as well as the posture of a person working out, causing key ROIs to appear and disappear. This also led to tracking errors. In future work, the concept of wearable thermal imaging used in Chapter 5 will be adopted in order to deal with the issue, and more than 20 participants will be recruited in order to reach statistically significant results.

Appendix B. Extraction of Cardiac Pulse Rate

In this section, we briefly show how to recover a person's cardiac pulse rate from thermal images. Using the proposed automatic ROI tracking algorithm presented in Chapters 4 and 5, regional thermal sequences focusing on a blood vessels shown on the neck (see Figure B.1) were extracted, followed with maximum thermal signals which were collected and processed with a series of the proposed method including interpolation, a high-pass filter (with cutoff frequency of 0.5 Hz; 30 beat-per-minute) and fast Fourier transformation (with a 7.5 seconds sliding window). To be specific, a cardiac pulse frequency is determined by searching local maxima between the average range of a healthy adult's heart rate (60 bpm – 120 bpm) (Karjalainen *et al.*, 1994; Ropers *et al.*, 2003). Furthermore, a thresholding technique is applied to remove ambiguous data points.

Figure B.1b shows the calculated heart rates (empty circles) using this method and heart rate signals (filled circles) simultaneously collected from the wearable PPG sensor (Figure A.3) as a ground truth. Overall, these results show a similar tendency in the cardiac rates from the selected 120s sequences. Here, sparse points in the heart rate data collected from the PPG sensor (see Figure B.1b, filled circles) were due to calculation faults in relation to motion or contact artefacts.

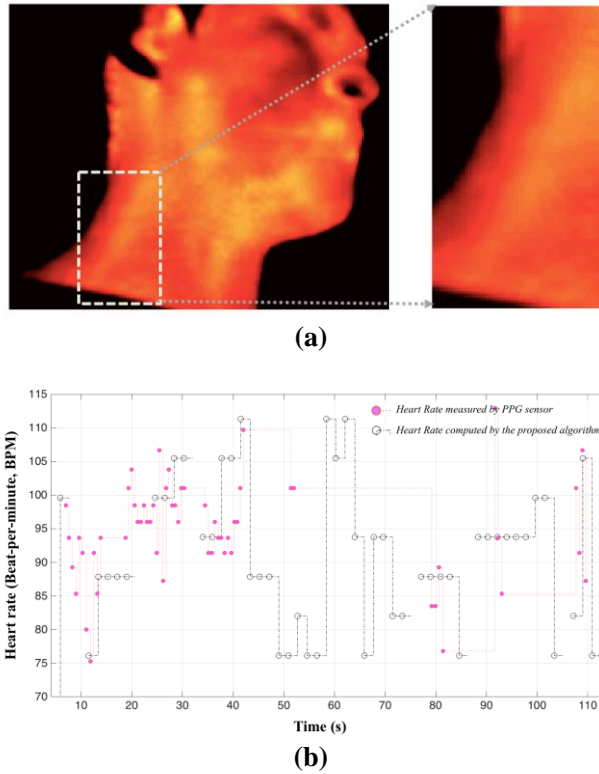


Figure B.1. (a) Blood vessel on the neck, (b) results: cardiac pulse rate measured by the proposed method (Grey-empty circles) and the rate measured by PPG sensor (Pink-filled circles).

Furthermore, using the *ThermalBreath II* dataset collected in Chapter 5, we evaluated performance of mobile thermal imaging-based pulse rate extraction using Pearson's correlation coefficients and additionally, the customised accuracy index CAND used in Garbey *et al.* (2007) to align the analysis process, which can be expressed as:

$$\text{CAND-base accuracy (in Garbey et al. 2007)} = 1 - \left| \frac{\mathcal{R}_{ref} - \mathcal{R}_m}{\mathcal{R}_{ref}} \right| \quad (\text{B.1})$$

where \mathcal{R}_{ref} is a value collected from references and \mathcal{R}_m is a physiological rate estimated through mobile thermal imaging (here, we call this approach, *ThermSense*).

Figure B.2 shows the results (correlation: $r=0.519$, $p<0.0001$, **CAND accuracy=90.79%**, $SD=7.364$). It shows the great potential in monitoring this physiological metric using mobile thermal imaging under uncontrolled situations including extreme conditions, such as sudden appearances of hot and cold temperatures in the background. However, as discussed in Section 2.2, we identify that the bespoke accuracy evaluation tool (**CAND**) tends to generally produce high values even in the case where correlation is weak. Given the weak correlation, there is a need to improve this performance so that we can use thermal imaging in measuring cardiac pulse rates.

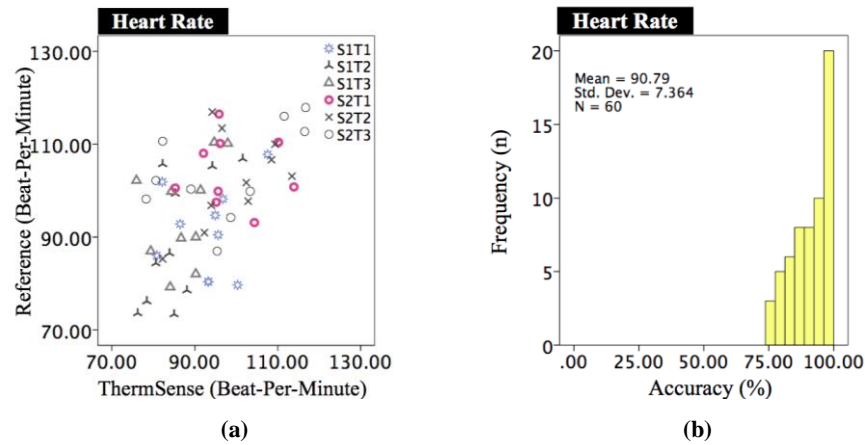


Figure B.2. (a) Scatter plot and (b) the CAND-based accuracy histogram of cardiac pulse rates extracted through mobile thermal imaging in comparison with reference data (Infiniti finger BVP sensor).

Appendix C. Extraction of Eye Blinks

Although little has been known about how to extract eye blinks from thermal image sequences, it is an important physiological response towards a person's psychological states (Caffier *et al.*, 2003). However, it is difficult to automatically extract one's eye-blink pattern from thermal images collected by low resolution (i.e. temporally and spatially), mobile thermal imaging since areas around one's eye are likely to get blurred in a thermal image of low spatial resolution. To recognise thermal changes on a ROI (i.e. eye) due to eye blinks, we propose the use of *Skewness* among features introduced in Section G. It can assist with sensing a moment when an eye blinks, as can be seen in Figure C.1. Figure C.1a illustrates the eye-ROI on thermal image sequences during the moments (collected at 186.545s, 186.713s and 186.869s) and when local peaks of *Skewness* are matched with the blinks, as shown in Figure C.1b. In future work, this approach will be evaluated into greater depth.

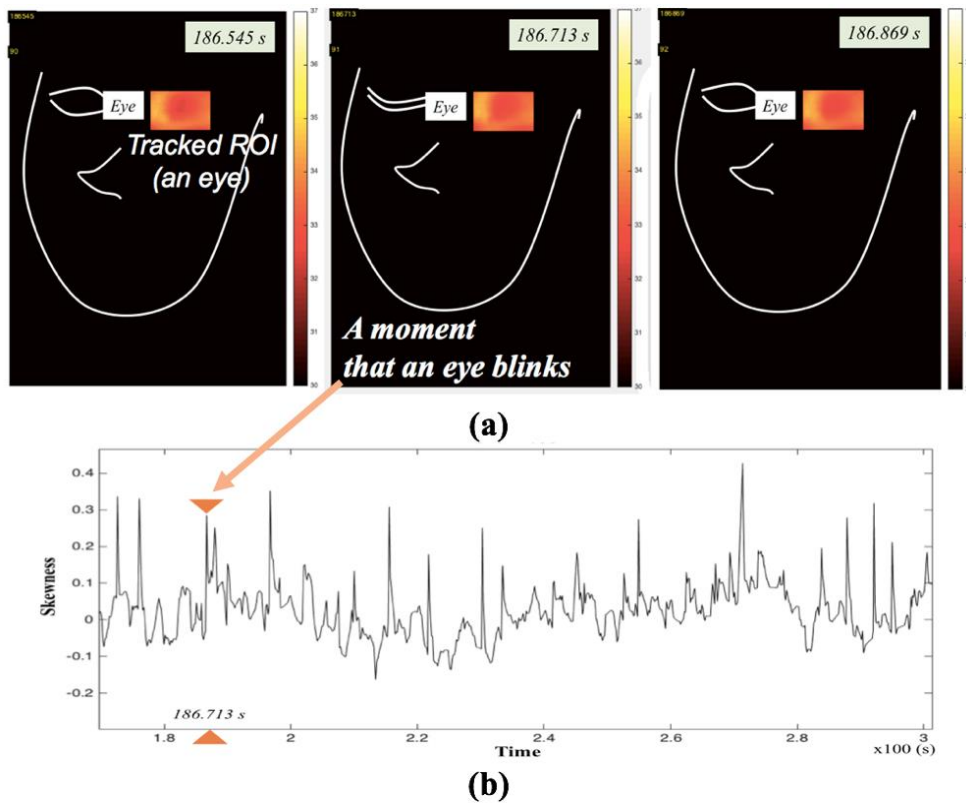


Figure C.1. Extraction of eye blinks: (a) the eye-ROI on thermal image sequences during the eye-blink moment, (b) skewness along with time.

Appendix D. Exploration of Non-linearity in Quantisation Process

D.1 Method

Non-linearity can be taken into consideration for the quantisation process in case a thermal dynamic range of a certain ROI is predictable, for example, thermal changes depending only on homoeothermic blood flows through a vessel in controlled ambient temperature for further enhancing tracking performance. By contrast with a linear quantisation (Figure D.1a), a non-linear quantisation can lessen colour variances caused by thermal dynamic changes (Figure D.1b,c). It can be expressed as

$$f(T) = \begin{cases} 0 & , \quad T < T_{min} \\ a_1(T - T_{mc})^{q/2} + a_2, & T_{min} \leq T \leq T_{max} \\ 2^n - 1 & , \quad T > T_{max} \end{cases} \quad (D.1)$$

$$f(T_{max}) = 2^n - 1, \quad f(T_{min}) = 0 \quad (D.2)$$

where T is temperature, T_{min} is a minimum temperature of interest, T_{max} is a maximum temperature of interest, and T_{mc} is an average temperature on a targeted ROI, n is the number of bits (e.g. 8bits-pixel or 16bits-pixels), a_1 and a_2 are coefficients which can be calculated by Equation (D.2). q is another coefficient which can be chosen by the predictable range of temperature variance on the targeted ROI. Relying on the coefficient q in Equation (D.1), the non-linear quantisation decides the size of a flat section, minimizing effects of thermal changes upon colour variations. In the case of regions showing high temperature changes (e.g. thermal variances in the nostrils by the respiration activity), q should be chosen to make the flat section wide. The coefficient q can be selected either heuristically or statistically.

Figure D.2 illustrates the concept of nonlinear quantisation. It can replace the optimal quantisation stage for the overall automated ROI tracking procedure (Figure 4.2). This could produce faster performances than the proposed optimal quantisation as it does not require the iterative computation used in our optimal quantisation. Nonetheless, these proposed methods can be complementary to each other since the maximum and minimum temperature of interest can be set to output data from the proposed optimal quantisation.

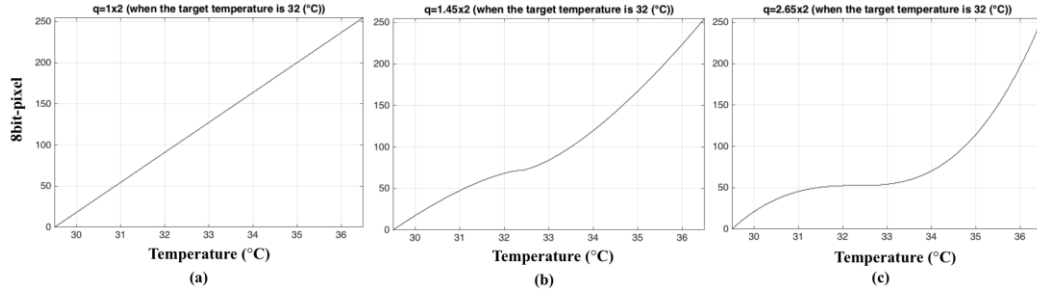


Figure D.1. Quantisation graphs: (a) $q=1.00 \times 2$, (b) $q=1.65 \times 2$, (c) $q=2.65 \times 2$

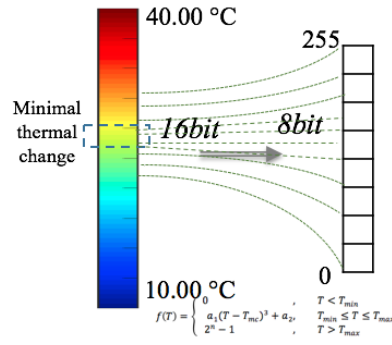


Figure D.2. A concept of Non-linear Quantisation

D.2. Results

To evaluate this concept, we implemented the Mode-seeking algorithm as a main tracker that was different from previous set ups shown in Section 4.3. This section showed the worst performance in the previous evaluation. This was to see how this non-linear quantisation concept can improve worst case scenarios. The Mode-seeking method generates a confidence map of the next frame from an image based on the histogram information (Cheng, 1995), and our non-linear Quantisation technique which can be applied before creating each confidence map corresponding to every frame sequence.

Without the use of the experimental protocols proposed in Section 3.3, thermal images of several human bodies were recorded using both a high-precision thermal camera and a low cost mobile thermal camera (see Figure 1.1) for this evaluation (only for collecting data from different body areas). Examples of the collected data (see Figure D.3) show different quantised thermal images of a human clenched hand along with different values of q in Equation (D.1). It is

noteworthy that the non-linear quantisation can smoothen colour differentials over the boundary of a targeted area (e.g. blood vessels).

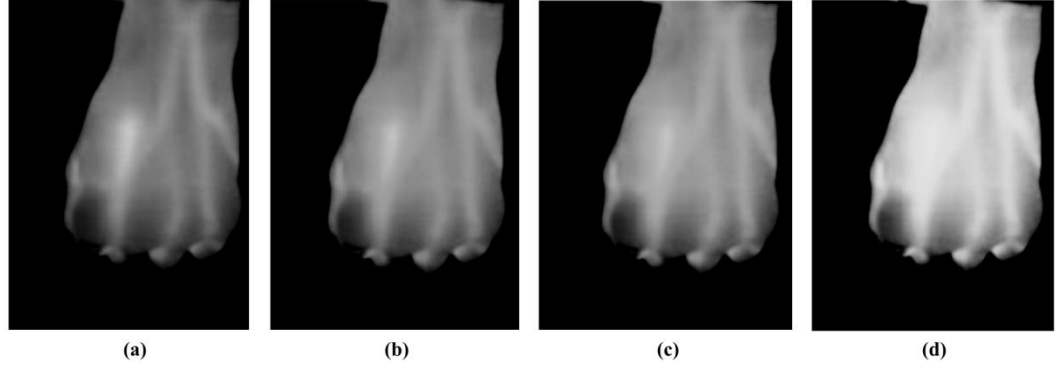


Figure D.3. Quantisation results of an image (a clenched hand and blood vessel, **target area: blood vessel**): (a) $q=1.00 \times 2$, (b) $q=1.30 \times 2$, (c) $q=1.45 \times 2$, (d) $q=1.65 \times 2$

Figure D.4 shows the results from the mode-seeking method with the nonlinear quantisation. As can be seen, one side of a parting of hair was selected as a ROI and the bounding box tracked the ROI completely. Table D.1 summarises the overall results achieved on different scenes, areas and ROIs. Here, five different scenes were recorded and eight ROIs were selected from the scenes (in this initial evaluation, thermal video samples collected from participants without protocols were used and tested). The performance of the tracking method with and without the pre-processing technique were compared. The coefficient q was chosen differently in accordance with the expected thermal variation range of each ROI. Overall, the tracking with the nonlinear Quantisation over-performed in comparison with the linear one.

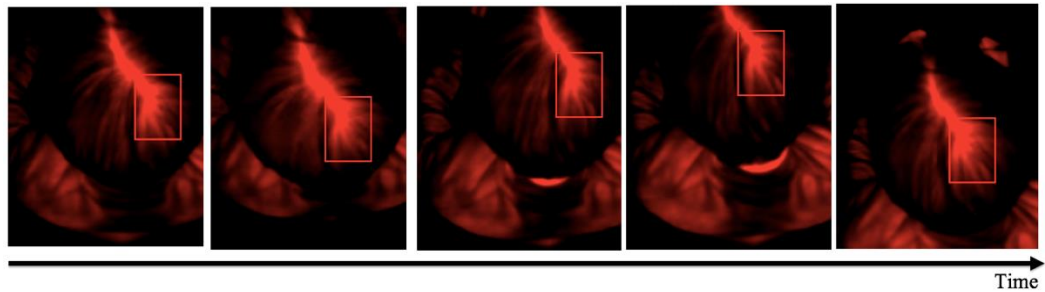


Figure D.4. Results from the mode-seeking with the proposed non-linear quantisation

Table D.1. Tracking results (based on Mean shift algorithm (Cheng, 1995)): number of successfully tracked frames without vs. with nonlinear quantisation ($q=1.45 \times 2$, 1.65×2 , 2.65×2)

Recorded area	Region of Interest	Total frame	Without Nonlinear Quantisation	With Nonlinear Quantisation
Man's head (posterior-view)	1) Whole head	4385 frames	4385 frames	4385 frames ($q=1.45 \times 2$)
	2) Parting		1530 frames	4385 frames ($q=1.45 \times 2$)
Man's head (top-view)	3) Parting	1667 frames	1667 frames	1667 frames ($q=1.45 \times 2$)
	4) Side scalp		1010 frames	1632 frames ($q=1.30 \times 2$)
Man's hand	5) Blood vessel	629 frames	163 frames	452 frames ($q=1.65 \times 2$)
Woman's face (anterior-view)	6) Nose/Nostril	3792 frames	1189 (1 st Failure) / 3096 (2 nd Failure) frames	3792 frames ($q=2.65 \times 2$)
	7) Eye		1106 frames	1339 frames ($q=1.45 \times 2$)
Man's neck (lateral-view)	8) Blood vessel	3871 frames	352 frames	2236 frames ($q=1.45 \times 2$)

Appendix E. Additional Materials for Deep Thermal Imaging (Material Type Recognition)

To compare the thermal dynamics levels of outdoor materials in Section 4.6 with those of indoor materials which can be found in a controlled laboratory or indoor places, we collected more thermal images of material surfaces from indoor materials, as illustrated in Figure E.1. Indoor materials are encountered in the context of a house or an office where the environment is typically more stable and controlled. The selection of indoor materials was mainly based on recent works (Sato *et al.*, 2015; Yeo *et al.*, 2017), however this was extended to include new classes in order to understand the limits of the approach (e.g., bubble wrap and paper towels showing similarity with freezer polyethylene bags and paper). Figure E.2 shows the thermal dynamic ranges of images collected from 15 indoor materials. The thermal dynamics ranged between 0.259°C and 3.507°C (denoted as A in Figure 7) with a mean of 0.889°C ($\text{SD}=0.4608$).

Example Materials in the Wild			
Indoor Materials			
a	<i>paper surface</i>	i	<i>hand palm</i>
b	<i>MDF material</i>	j	<i>wood surface</i>
c	<i>MFC table tops</i>	k	<i>calfskin</i>
d	<i>nylon stiff fabric</i>	l	<i>fur fabric</i>
e	<i>plastic freezer bag</i>	m	<i>card board</i>
f	<i>PVC leather</i>	n	<i>bubble wrap</i>
g	<i>paper towel</i>	o	<i>TPE material</i>
h	<i>concrete wall covered by matt emulsion</i>		

Figure E.1. Example materials collected in the wild: (a) materials generally found in indoor places (less influenced by environmental thermal dynamics),

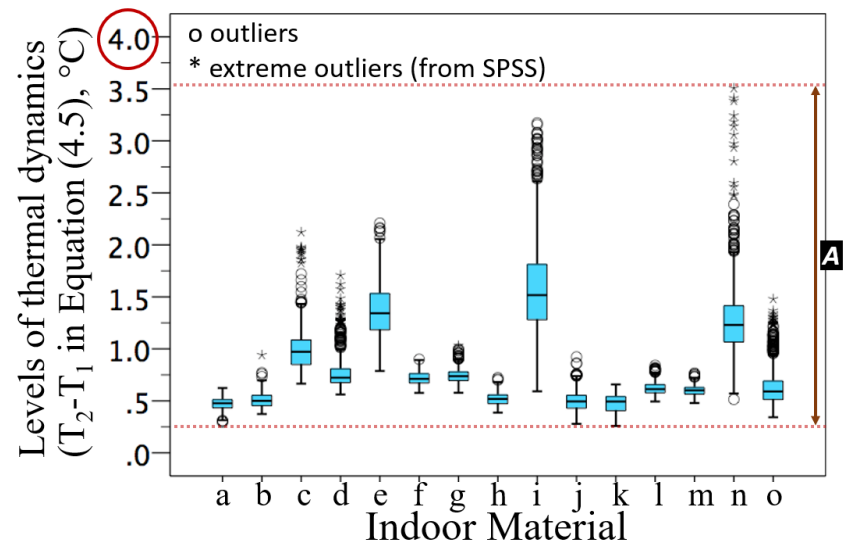


Figure E.2. Levels of thermal dynamics of the total 14860 thermal images across the 15 indoor materials (a-o). The thermal dynamics were estimated by using ($T_2 - T_1$) in Equation (4.5). [A] indicates the range of those values.

Appendix F. BVP and P-P Interval Estimation through PPG Imaging

As part of the methods proposed for building the cardiovascular patterns-based instant stress recogniser in Chapter 9, we additionally propose a technique to produce reliable PPG-derived HRV profiles from a smartphone camera²².

F.1. Spatial Entropy-based camera PPG

Figure F.1 summarises the approach we use to extract BVP and P-P intervals through the smartphone imaging PPG. Following existing methods (Jonathan & Leahy, 2010; Chan *et al.*, 2016; White & Flaker, 2017), our method estimates the BVP signals by capturing subtle colour variations associated with light absorptivity patterns of haemoglobin in the capillaries of a person's skin. However, rather than using average values of the pixels of the red (or green) channel to estimate the BVP value, which is the most widely used method (Jonathan & Leahy, 2010; McManus *et al.*, 2013; Chan *et al.*, 2016), we propose to use the temporal variations in spatial Shannon's entropy (Shannon, 1948) of sequential R-channel images as raw BVP signals. This is due to averaging which tends to ignore fairly small but important variations in colour distribution (as in Chapter 5). The estimated BVP value $B_t(X)$ at a given time t can be expressed as:

$$B_t(X) \equiv -H_t(X) = \sum_{(i,j)} p(x_{i,j}) \log_2 p(x_{i,j}) \quad (\text{F.1})$$

where $x_{i,j}$ is the brightness of pixel(i,j), $H_t(X)$ is the entropy function and $p(x_{i,j})$ is the probability distribution which is generally estimated using a grayscale histogram in image analysis (Sonka *et al.*, 2014) (here, for the R channel).

²² Details can be found from Cho *et al.* (2019).

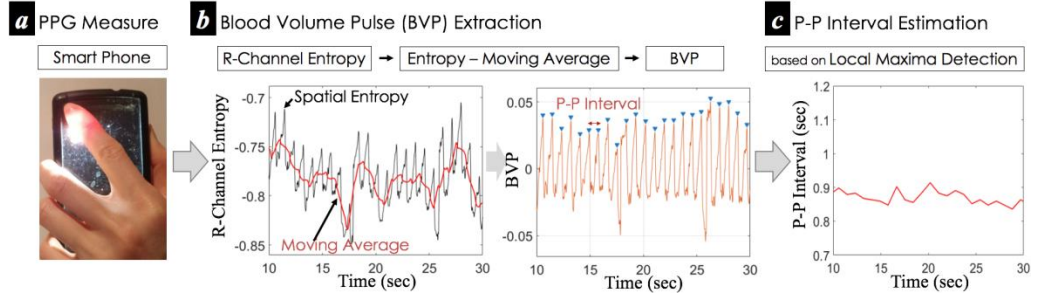


Figure F.1. Overall procedure of BVP and P-P interval estimation from a person's finger through the smartphone-imaging PPG. See text for details.

As our interest is in measuring raw P-P intervals from PPG signals, we used a simple signal processing technique to create similar amplitudes of each peak of BVP which helps detect peaks for measuring the time interval (i.e. P-P interval) between the peaks. This was calculated by the subtraction of the k -sample moving average signals from the raw entropy signal (Figure F.1b), which can be expressed by:

$$\hat{B}_t = B_t - \frac{1}{k} \sum_{i=t-k}^{t-1} B_i. \quad (\text{F.2})$$

As a high sampling rate produces a higher sensitivity of the P-P intervals (Kumar et al., 2015), we up-sampled the raw sequences to 256 Hz with spline interpolation and used a 1s moving average to smooth heartbeat induced variations within the duration where at least one heartbeat of a normal person is expected to appear (Tsuji et al., 1996). Finally, we used the simple local maxima detection (Hamilton & Tompkins, 1986) with a 0.5 second sliding window to recover P-P intervals (Figure F.1c).

F.2. Result of BVP Signal Quality Test

As the cardiac measurement capability of smartphone PPG has previously been thoroughly investigated in earlier studies (McManus et al., 2013; Chan et al., 2016; White & Flaker, 2017), we only tested the reliability of the cardiac pulse signals measured with our approach by comparing it with signals from the widely-used approach of mean brightness intensity (Jonathan & Leahy, 2010; McManus et al., 2013; Chan et al., 2016). For this, we used the relative power Signal Quality Index (pSQI) as seen in Chapter 6 (Section 6.4, Equation 6.4), as a measure of

quality (Clifford *et al.*, 2012; Kumar *et al.*, 2015; Elgendi, 2016). The pSQI for the BVP signals can be expressed by:

$$P(\hat{f}_{\min} \leq f \leq \hat{f}_{\max}) \cong \frac{\int_{\hat{f}_{\min}}^{\hat{f}_{\max}} S_{\hat{B}}(f) df}{\int_{total} S_{\hat{B}}(f) df} \quad (F.3)$$

where $0 \leq P \leq 1$, $S_{\hat{B}}$ is the power spectral density of BVP signals (in our case, \hat{B} in Equation (F.2)), and \hat{f}_{\min} , \hat{f}_{\max} are the lower and upper boundary of expected HRs, respectively. Here, we set the expected HR range to [0.8Hz (48bpm), 2.0Hz (120bpm)] given that HRs of healthy adults mostly fall into this range (Tsuji *et al.*, 1996). To minimize effects of the baseline wander and high-frequency noise on this signal quality test (Chan *et al.*, 2016; Elgendi, 2016), we used band-pass filtered BVP signals (0.7-4.0Hz, Chan *et al.*, 2016). Figure F.2 compares the pSQI values of the proposed method (Equation (F.2)) with those of the mean intensity method, showing improved quality of the estimated BVP signals \hat{B} (Proposed: M=0.755, SD=0.068; Traditional: M=0.692, SD=0.075).

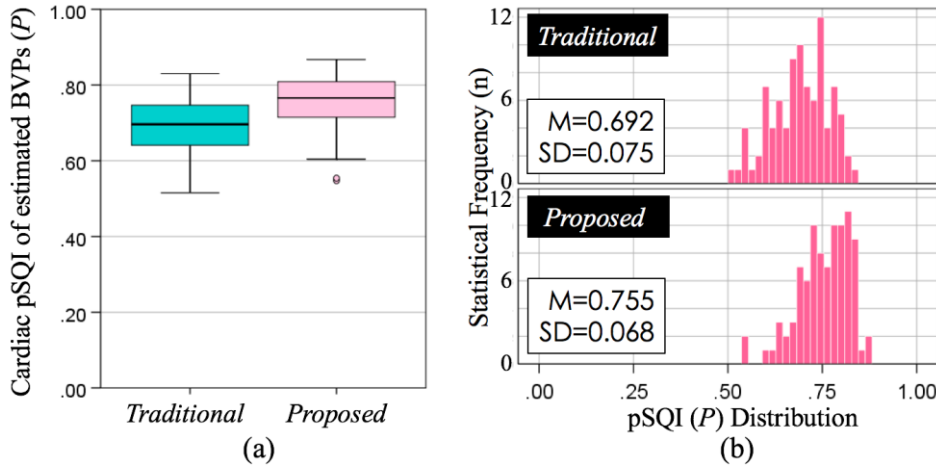


Figure F.2. Signal extraction quality comparison of our spatial entropy-based method (Equation(2)) with the mean intensity approach (Jonathan & Leahy, 2010; McManus *et al.*, 2013; Chan *et al.*, 2016) by using pSQI: (a) box plot, (b) histogram.

Appendix G. Hand-Engineered Features for Thermal Analysis

Based on the findings from earlier studies on thermography-based affect recognition and machine learning topics, we classify general hand-engineered features into three categories: i) basic statistical features, ii) information-theoretic features, and iii) frequency features. Firstly, the basic statistical features range from standard deviation, skewness and kurtosis, to cross-correlation and integral ratio based on thresholding. Secondly, the information-theoretic features can be exemplified by Shannon's entropy (Shannon, 1948). Entropy and cross-entropy are examples of these features. Lastly, the frequency domain features include a peak frequency, a bandwidth and a wavelet. They can also be used as metrics for quantifying physiological signatures. Table G.1 summarises the features. The features can be computed from a specific ROI on thermal images.

Table G.1. Summary of feature candidates for thermal analysis.

Feature and abbreviation	Definition
<i>i) basic statistical features</i>	
Differential average temperature (ΔT_{av})	Amount of difference between two average temperatures on a ROI collected from specific moments.
Maximum temperature (T_{max})	Arithmetic maximum temperature on a two-dimensional thermogram.
Standard deviation of temperature (σ_T)	Quantification of the amount of dispersion of a two-dimensional temperature distribution on a ROI.
Skewness of temperature (γ_T)	A measure of the asymmetry for the probability distribution of a two-dimensional thermogram.
Kurtosis of temperature (K_T)	A measure of the tailedness for probability distribution of a two-dimensional thermogram.
Histogram of Oriented Gradients of Thermogram (HOG_T)	Histogram counting occurrences of gradient orientation in localized portions of temperature patterns on a ROI.
Two-dimensional cross correlation ($f_T * g_T$)	A measure of similarity of two dimensional thermograms.
Integral ratio based on thresholding (IR_{th})	Ratio of integral thermal values over a specific threshold to the total size of a ROI.
<i>ii) information-theoretic features</i>	
Entropy (S)	A measure of unpredictability of a thermal pattern on a ROI.
Cross entropy of thermograms (CE_T)	A measure of the difference between two thermal distributions.
<i>iii) frequency features</i>	
Peak frequency (f_{peak})	Arithmetic maximum frequency in a power spectrum.
Bandwidth (B_f)	Width of valid frequency bands in a power spectrum.
Wavelet based time-frequency band (WB_{t-f})	A measure of the relative energy in a time-frequency domain.

Appendix H. Questionnaires

Demographics

1. Participant Number

2. Gender ☐ Male ☐ Female

3. Age

[Optional] ethnic group?

4. Job ☐ Student ☐ Researcher ☐ Faculty ☐ School Staff ☐ Employed in a company ☐ Other

5. Do you have any history of the following..

☐ neurological disorder ☐ substance abuse ☐ severe physical illness

☐ psychiatric disorder ☐ head injury

☐ Any others? []

6. Please choose one response from the four given for each item.

a. I feel mentally stressed (ordinarily)

☐ Not at all ☐ Not often ☐ Sometimes ☐ Most of the time

b. I feel physically tired I (ordinarily)

☐ Not at all ☐ Not often ☐ Sometimes ☐ Most of the time

7. Have you experienced this type of task before? []

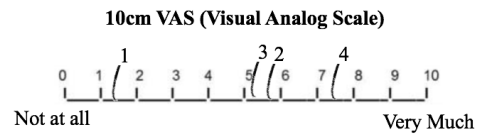
If so, when was the last time? []

How many times? []

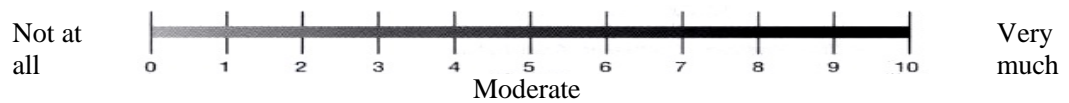
Main questionnaire used in Stress-induction Tasks (Page 1)

You will be asked to rate your subjective scores along with each component.
Please rate each score with its own session number (e.g. 1,2,3 ...). For example:

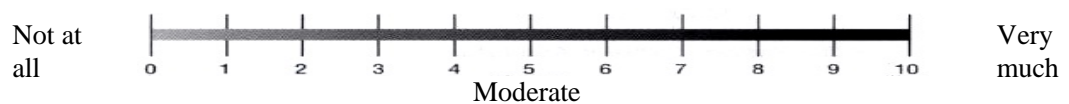
e.g.) “How much did you feel *mentally* stressed?”



Q1) How much mental stress did you experience?



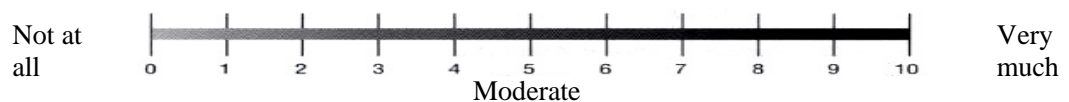
Q2) How much physical tiredness did you experience?



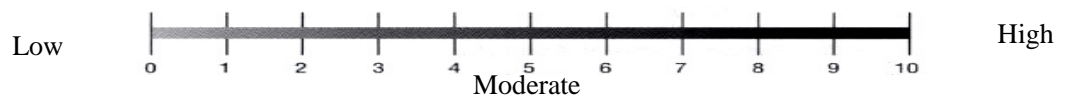
Q3) How much cognitive load did you experience?



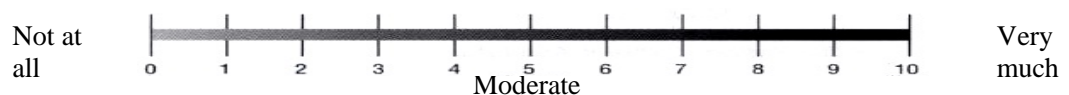
Q4) How relaxed did you feel?



Q5) Difficulty of the task?



Q6) How engaged did you feel with the task?



Main questionnaire used in Stress-induction Tasks (Page 2~)
(Repeated question sets)

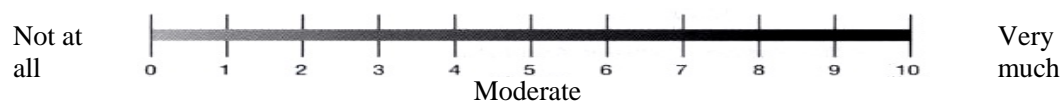
Q1) How much mental stress did you experience?



Q2) How much physical tiredness did you experience?



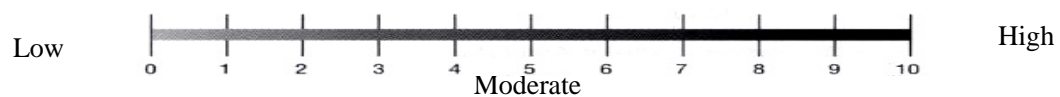
Q3) How much cognitive load did you experience?



Q4) How relaxed did you feel?



Q5) Difficulty of the task?



Q6) How engaged did you feel with the task?



Appendix I. Ethical Approval

Information Sheet I

You will be given a copy of this information sheet.

Title of Project: **Emotion & Pain Project**

This study has been approved by the UCL
Research Ethics Committee [Project ID
Number]: STAFF/1011/005

Name, Address and Contact Details of
Investigators:

Prof. Nadia Bianchi-Berthouze
University College London Interaction
Centre
2nd Floor, 66-72 Gower Street
London WC1E 6EA, United Kingdom
+44 (0)20 3108 7067

We would like to invite you to participate in this research project. You should only participate if you want to; choosing not to take part will not disadvantage you in any way. Before you decide whether you want to take part, it is important for you to read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is not clear or if you would like more information.

This project aims to develop technology to help patients with chronic pain by providing tailored feedback and support for movements performed as part of self-directed rehabilitation.

We may interview you about the needs and uses for such technology. We will ask your opinion of current prototypes we have developed. We will ask you to do everyday activities, exercise or play computer games whilst wearing movement sensors and/or biosensors; The activities will be recorded using these sensors, thermal cameras, and video/audio recording. We will also ask you to complete pain questions or movement related questionnaires.

All data will be handled according to the Data Protection Act 1998 and will be kept anonymous. Researchers working with Prof. Berthouze will analyse the data collected. The information gathered will be used to understand requirements of technology intervention for mental health and pain management.

With your permission, we would like to use extracts of the video and audio recordings to demonstrate to people how assistive technology can be used for the management of their condition.

With your permission, we would also like to use extracts of the video and audio recordings for teaching, conferences, presentations, publications, and/or thesis work. Please note that these presentations may be recorded by individuals without our knowledge and displayed on social media.

It is up to you to decide whether or not to take part. If you choose not to participate, it will involve no penalty or loss of benefits to which you are otherwise entitled. If you decide to take part, you will be given this information sheet to keep and be asked to sign a consent form. If you decide to take part, you are still free to withdraw at any time without providing a reason.

Informed Consent Form I

(This form is to be completed independently by the participant after reading the Information Sheet and/or having listened to an explanation about the research.)

Title of Project: **Emotion & Pain Project**

This study has been approved by the UCL Research Ethics Committee [Project ID Number]: STAFF/1011/005

Participant's Statement

I

agree that I have

- read the information sheet and/or the project has been explained to me orally;
- had the opportunity to ask questions and discuss the study;
- received satisfactory answers to all my questions or have been advised of an individual to contact for answers to pertinent questions about the research, my rights as a participant and whom to contact in the event of a research-related injury.

I understand that my participation will be taped/video/sensors recorded and I am aware of and consent to the analysis of the recordings;

I understand that I must not take part if I am not physically able to do the tasks;

I agree to be invited in the future by UCL researchers to participate in follow-up studies.

I agree for the videotape to be used by the researchers in this project and in further research studies.

I agree for the videotape to be used by the researchers to demonstrate assistive technology to clinicians and people with chronic pain. and clinicians.

I agree for the videotape to be used by the researchers for teaching, conferences, presentations, publications, and/or thesis work.

I understand that these presentations may be recorded without the knowledge of the researchers by via media or other individuals/researchers.

I agree for the videotape to be used in other projects by members of this research group.

I agree for the videotape to be open to the community inside/outside UCL

I understand that I am free to withdraw from the study without penalty if I so wish and I consent to the processing of my personal information for the purposes of this study only and that it will not be used for any other purpose. I understand that such information will be treated as strictly confidential and handled in accordance with the provisions of the Data Protection Act 1998.

Signed:

Date:

Investigator's Statement

I

confirm that I have carefully explained the purpose of the study to the participant and outlined any reasonably foreseeable risks or benefits (where applicable).

Signed:

Date:

Information Sheet II

You will be given a copy of this information sheet.

Title of Project: **HUMAN MANufacturing (HUMAN)**

This study has been approved by the UCL Research Ethics Committee [Project ID Number]: UCLIC/1617/003/Staff

Julier/Berthouze

Name, Address and Contact Details of Principal Investigator:

Dr. Simon Julier
University College London Interaction Centre
5th Floor, 66-72 Gower Street
London WC1E 6EA, United Kingdom
+44 (0)20 3108 7114

We would like to invite you to participate in the EU H2020 HUMAN research project (<http://www.humanmanufacturing.eu/>). You should only participate if you want to and feel capable of doing so; choosing not to take part will not disadvantage you in any way. Before you decide whether you want to take part, it is important for you to read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is not clear or if you would like more information.

The project will develop a solution that combines factory-level sensing systems (which monitor current status and progress) with assistance systems to provide cognitive and physical support to help workers when conducting difficult and challenging tasks. We will ask you to conduct mental stress induction tasks which are shown on a screen whilst using a chin rest to limit your head movement. We may also ask you to work at a quicker pace to explore the effects of time-related stress.

We will ask for your permission to record you using thermal cameras and neural/physiological sensing data capture devices. We are investigating whether such sensing systems could be used to identify your stress levels and affective states. If you were to feel uncomfortable with any of the sensors or cognitive-load induction tasks you should inform us and we will stop and remove the sensors at once. You will also be able to take a break any time you need.

In the case of thermal cameras, you may be visible but hardly recognizable. All the thermal cameras present will be shown and explained to you so that you can make a decision as to whether you wish to allow the recording or not.. You can also ask to look at the thermal video at the end of the study and/or to delete this video if you feel uncomfortable about what you did during the study. For the sensor data, we record information about blood volume pulse, breathing, skin conductance, muscle activation patterns, cortical activity, and movement. All the sensors present will be shown and explained to you so that you can make an informed decision as to whether

or not to allow the recording. All the sensor data will be anonymously recorded. We will analyse and process the anonymous thermal video and data for building a machine learning-based system that could recognize your affective states.

Again, with your permission, we would like to use extracts of the thermal video and sensor data recordings for teaching, conferences, presentations, publications, and/or thesis work. Please note that these presentations may be recorded by individuals without our knowledge and displayed on social media. You will be allowed to decide if you grant us the permission to show the thermal videos even though the person is hardly identifiable on every thermal video frame.

All data will be handled according to the Data Protection Act 1998 and will be kept anonymous. Researchers working with Dr. Julier will analyse the data collected. The information gathered will be used to understand how the developed technology can be used to assist work practices.

It is up to you to decide whether or not to take part. If you choose not to participate, there will be no penalty. If you decide to take part, you will be given this information sheet to keep and be asked to sign a consent form. If you decide to take part, you are still free to withdraw at any time during the study and without providing a reason.

* The data gathered by the Empatica bracelet (galvanic skin changes, heart rate and blood volume pulse, temperature, acceleration of the arm) are automatically stored in the EMPATICA Ltd. cloud outside of the UK. The data are associated to a number that has no relation with the person and will be fully anonymous.

Informed Consent Form II

(This form is to be completed independently by the participant after reading the Information Sheet and/or having listened to an explanation about the research.)

Title of Project: **HUMAN MANufacturing (HUMAN)**

This study has been approved by the UCL Research Ethics Committee [Project ID Number]: UCLIC/1617/003/Staff Julier/Berthouze

NOTE: as you agree to the study, please circle and delete as appropriate multiple choices listed between "[...]".

Participant's Statement

I

agree that I have

- read the information sheet and/or the project has been explained to me orally;
- had the opportunity to ask questions and discuss the study;
- received satisfactory answers to all my questions or have been advised of an individual to contact for answers to pertinent questions about the

research and my rights as a participant and whom to contact in the event of a research-related injury;

- I understand and consent for the experiment which can be run by any member of the HUMAN project (and not just UCL), however I also understand that UCL will handle the data storage according to the agreement set below;
- I understand that the data gathered by the Empatica bracelet (galvanic skin changes, heart rate and blood volume pulse, temperature, acceleration of the arm) are automatically stored in the EMPATICA Ltd. cloud outside of the UK. The data are associated to a number that has no relation with the person, and will be fully anonymous.

For the following, mark each point and circle as appropriate for choices in “[...]”:

I **consent** for the researcher to take thermal video recording of me as I participate in the experiment..;

I **consent** for the researcher to wear sensors, the head mounted display and record sensor data as I participate in the experiment. I understand that I can ask to remove any of the sensors and the chin rest if I feel uncomfortable;

I **consent** for the anonymous **[thermal video / sensor data / notes]** recording to be published in academic publications, presented at conferences and in lectures or through public engagement activity.

I **consent** for the anonymous **[thermal video / sensor data / notes]** recording to be disclosed as part of a data set to foster research in research communities.

PERMISSION OF USE OF NON ANONYMIZED THERMAL VIDEO DATA

For the following, initial each point and circle as appropriate for choices in “[...]”:

____ I agree for the thermal videotape to be used by UCL researchers in this project and in further research studies **[YES / NO]**

____ I agree for the thermal videotape to be used by UCL researchers for teaching, conferences, presentations, publications, and/or thesis works. I understand that these presentations may be recorded without the knowledge of the researchers via media or other individuals / researchers industries

[YES / NO]

____ I agree for the thermal videotape to be disclosed as part of the data set to foster research in research communities.

[YES / NO]

PERMISSION OF USE OF NON ANONYMIZED SENSOR DATA (i.e. non anonymized transcripts)

For the following, mark each point and circle as appropriate for choices in “[...]”.

____ I agree for the sensor data recording to be used by UCL researchers in this project and in further research studies **[YES / NO]**

____ I agree for the sensor data recording to be used by UCL researchers for teaching, conferences, presentations, publications, and/or thesis works. I

understand that these presentations may be recorded without the knowledge of the researchers via media or other individuals/researchers industries

[YES / NO]

____ I agree for the sensor data recording to be disclosed as part of a data set to foster research in research communities.

[YES / NO]

Finally,

I agree to be invited in the future by UCL researchers to participate in follow-up studies.

I understand that I am free to withdraw from the study without a penalty if I so wish and I consent to the processing of my personal information for the purposes of this study only and that it will not be used for any other purpose. I understand that such information will be treated as strictly confidential and handled in accordance with the provisions of the Data Protection Act 1998.

Signed:

Date:

Investigator's Statement

I

confirm that I have carefully explained the purpose of the study to the participant and outlined any reasonably foreseeable risks or benefits (where applicable).

Signed:

Date: